MULTILEVEL PRECONDITIONERS FOR DISCONTINUOUS GALERKIN APPROXIMATIONS OF ELLIPTIC PROBLEMS WITH JUMP COEFFICIENTS

By

Blanca Ayuso de Dios

Michael Holst

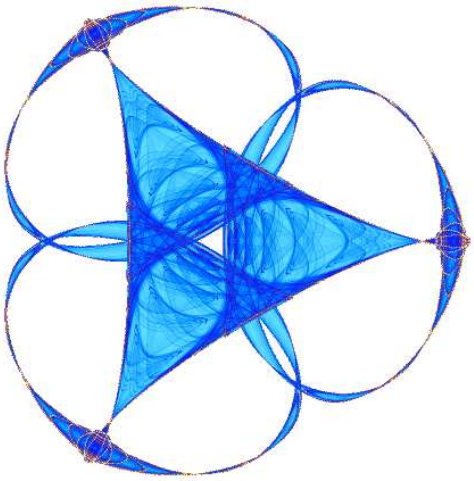
Yunrong Zhu

and

Ludmil T. Zikatanov

IMA Preprint Series # 2356

(December 2010)



INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS

UNIVERSITY OF MINNESOTA

400 Lind Hall

207 Church Street S.E.

Minneapolis, Minnesota 55455–0436

Phone: 612-624-6066 Fax: 612-626-7370

URL: http://www.ima.umn.edu

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
DEC 2010		**00-00-2010 to 00-00-2010**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Multilevel Preconditioners for Discontinuous Galerkin Approximations of Elliptic Problems with Jump Coefficients	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Minnesota,Institute for Mathematics and Its Applications,207 Church Street SE,Minneapolis,MN,55455-0436**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

In this article we develop and analyze two-level and multi-level methods for the family of Interior Penalty (IP) discontinuous Galerkin (DG) discretizations of second order elliptic problems with rough coe cients (exhibiting large jumps across interfaces in the domain). These methods are based on a decomposition of the DG nite element space that inherently hinges on the di usion coe cient of the problem. Our analysis of the proposed preconditioners is presented for both symmetric and non-symmetric IP schemes, and we establish both robustness with respect to the jump in the coe cient and near-optimality with respect to the mesh size. Following the analysis, we present a sequence of detailed numerical results which verify the theory and illustrate the performance of the methods.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**44**	

MULTILEVEL PRECONDITIONERS FOR DISCONTINUOUS GALERKIN APPROXIMATIONS OF ELLIPTIC PROBLEMS WITH JUMP COEFFICIENTS

BLANCA AYUSO DE DIOS, MICHAEL HOLST, YUNRONG ZHU, AND LUDMIL ZIKATANOV

Abstract. In this article we develop and analyze two-level and multi-level methods for the family of Interior Penalty (IP) discontinuous Galerkin (DG) discretizations of second order elliptic problems with rough coefficients (exhibiting large jumps across interfaces in the domain). These methods are based on a decomposition of the DG finite element space that inherently hinges on the diffusion coefficient of the problem. Our analysis of the proposed preconditioners is presented for both symmetric and non-symmetric IP schemes, and we establish both robustness with respect to the jump in the coefficient and near-optimality with respect to the mesh size. Following the analysis, we present a sequence of detailed numerical results which verify the theory and illustrate the performance of the methods.

Contents

Date: December 7, 2010.

Key words and phrases. Multilevel preconditioner, discontinuous Galerkin methods, Crouzeix-Raviart finite elements.

The work of the first author was partially supported by the Spanish MEC under projects MTM2008-03541 and HI2008-0173.

The work of the second author was supported in part by NSF DMS-0715146, NSF DMS-0915220, and DTRA Award HDTRA-09-1-0036.

The work of the third author was supported in part by NSF DMS-0715146 and DTRA Award HDTRA-09-1-0036.

The work of the fourth author was supported in part by NSF DMS-0810982 and NSF OCI-0749202.

1. Introduction

In this paper, we develop and analyze robust multilevel preconditioners for discontinuous Galerkin (DG) discretizations of the second order elliptic equation with strongly discontinuous coefficients:

$$\begin{cases} -\nabla \cdot (\kappa \nabla u) = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases} \tag{1.1}$$

where $\Omega \subset \mathbb{R}^d$ is a bounded polygon (for $d = 2$) or polyhedron (for $d = 3$). The scalar function $\kappa = \kappa(x)$ denotes the diffusion coefficient which is assumed to be piecewise constant with respect to an initial non-overlapping subdomain partition of the domain Ω, denoted $\mathcal{T}_S = \{\Omega_m\}_{m=1}^M$ with $\cup_{m=1}^M \overline{\Omega}_m = \overline{\Omega}$ and $\overset{o}{\Omega}_m \cap \overset{o}{\Omega}_n = \emptyset$. Although the (polygonal or polyhedral) regions Ω_m , $m = 1 \dots M$ might have complicated geometry, we will always assume that there is an initial triangulation $\mathcal{T}_0$ such that $\kappa_T = \kappa(x)|_T$ is a constant for all $T \in \mathcal{T}_0$. Problem (1.1) belongs to the class of interface or transmission problems, which are relevant to many applications such as groundwater flow [54], electromagnetics [53], semiconductor device modeling [31, 59], and fuel cells [74, 76]. The coefficients in these applications might have large discontinuities across the interfaces between different regions with different material properties.

Finite element discretizations of (1.1) lead to linear systems with badly conditioned stiffness matrices; not only is the ill-conditioning with respect to the mesh size, but the condition numbers depend linearly on the largest jump in the coefficients. Much research has been devoted to developing efficient and robust preconditioners for use with iterative methods for conforming finite element discretizations of (1.1). The resulting preconditioners can be grouped into two basic classes of methods: non-overlapping and overlapping methods. Both classes of preconditioners have their advantages and disadvantages, and as a result it is desirable to have access to robust and efficient methods from both classes. Domain decomposition Balancing Neumann-Neumann [58], FETI-DP [55] and Bramble-Pasciak-Schatz preconditioners [12] belong to the first class of methods (the non-overlapping methods). They have been shown to be robust with respect to coefficient variations and mesh size (up to a logarithmic factor), in theory as well as practice, but only if special coarse spaces and coarse solvers (such as those based on discrete harmonic extensions [42, 58, 43, 65]) are constructed (see also a survey paper by Xu and Zou [79] for many techniques and references).

The class of overlapping methods encompasses, among others, overlapping Schwarz methods, geometric multigrid, and more general Multivel methods, as such the Bramble-Pasciak-Xu (BPX) preconditioner (see e.g. [13, 14, 77]). In practice, it has always been observed that all these methods, when used as preconditioners in conjugate gradient iteration, result in efficient algorithms that behave robustly with respect to the jump in the coefficients, independently of the problem dimension. However, the first analyses show robustness only in certain particular cases: *quasi-monotonicity* assumption on the coefficients [41]; space dimension $d \leq 2$ [10]; or when each subdomain Ω_m touches the Dirichlet boundary [75, 61]. In general, these theories predict a deterioration in the rate of convergence of multigrid and overlapping domain decomposition methods, with respect to both the coefficients and the mesh size. An improvement in the rate of convergence can be achieved by resorting to basis stabilization techniques [70], or by employing energy minimizing coarse spaces [73].

However, we note that there is a discrepancy between these theoretical results and the convergence rate actually observed in practice. This is due to the fact that the analysis approach used in earlier works was based mainly on the standard theory of CG (see for example [51, Theorem 9.4.12], or [63, Theorem 6.29]) which provides only bounds in the worst case scenario, and does not exploit the spectral structure of particular preconditioned system. The preconditioned system may actually have only a few small eigenvalues (depending on the coefficient distribution

and mesh size), while the other eigenvalues are bounded nearly uniformly. This was first observed by [48, 72] for the simple Jacobi preconditioner. By using a more sophisticated approach to CG theory (see [6, Section 13.2], [7]) and considering the distribution of the spectrum of the preconditioned system, it is possible to show that the small eigenvalues do not influence the (observed) *asymptotic convergence rate*. This approach has been pursued in [78, 80], where it is show that both standard multilevel and overlapping domain decomposition methods lead to nearly optimal preconditioners for CG algorithms. See also [26] for further extensions in the case of highly graded adaptive meshes.

While preconditioners for conforming discretizations have been considered by a number of researchers, less work has focused on the construction of preconditioners for non-conforming approximations of (1.1). In this article, we consider preconditioners when the family of Interior Penalty (IP) Discontinuous Galerkin (DG) methods are used to approximate (1.1). DG methods are designed to provide robust discretizations of partial differential equations of nearly any type, and can even handle equations whose type changes within the computational domain. They are naturally suited for multi-physics applications, and for problems with highly varying material properties, such as problem (1.1). The development of efficient solvers for DG discretizations has been pursued only in the last ten years; however, due to the growing importance of DG methods, there is now considerable interest in this area. Examples of overlapping preconditioners for non-conforming discretizations are found in [65, 64], where the analysis depends on the assumption that the coefficient κ is quasi-monotone. Since the algorithms we develop here contain the solution of systems corresponding to non-conforming discretizations of the model problem (1.1) as a part of the preconditioning step, the analysis presented here gives bounds on the convergence rate for two level and multilevel methods for the lowest order non-conforming Crouseix-Raviart finite element discretizations.

While classical approaches have been successfully extended for second order elliptic problems, discontinuous nature of the underlying finite element spaces has motivated the creation of new techniques for the design of solvers. Additive Schwarz methods (of overlapping and non-overlapping type) are considered and analyzed in [44, 38, 2, 3, 4, 11]. Multigrid methods are studied in [47, 20, 19, 18, 62, 30]. Two level methods are presented in [34, 22, 23]. More general multi-level methods based on algebraic techniques are considered in [57, 56]. However, except from several numerical experiments reported in [34, 33, 3, 35, 56], all these works deal with the case of a smoothly or slowly varying diffusivity coefficient. For problem (1.1), only in [38, 39, 40] the authors introduce and analyze non-overlapping BBDC and FETI-DP domain decomposition preconditioners for a Symmetric Interior Penalty discretization of (1.1). The DG discretization in only used on the skeleton of the subdomain partition, while a standard conforming approximation is used in the interior of the subdomains. Robustness and quasi-optimality is shown for the Additive and Hybrid BBDC [39] and FETI-DP [40] preconditioners, even for the case of non-matching grids. The analysis of the first preconditioner requires an *interface condition* relating the magnitude of the coefficient and the mesh-size.

The goal of this article is to design, and provide a rigorous analysis of, a simple multilevel solver (belonging to the overlapping category), for the family of Interior Penalty methods for approximating (1.1). Together with the family of IPDG methods (including symmetric and non-symmetric methods), we consider the corresponding family of *weakly penalized* or IPDG-0 methods (called Type-0 in [9]). Our approach follows the ideas in [9], and it is based on a splitting of the DG space into two components that are orthogonal in an energy inner product (more precisely, in the energy inner product induced by the IPDG-0 methods). Roughly speaking, the construction amounts to identifying a "low frequency" space (Crouzeix-Raviart elements) and then defining a second complementary space. Such a decomposition turns out to be critical to the design of robust preconditioners in the case of PDEs with rough coefficients, such as problem (1.1). However, a notable difference takes place in the decomposition of the DG space

introduced for the Laplace equation [24, 9]. For problem (1.1), a stable decomposition can be obtained by introducing subspaces that depend on the coefficient κ, and this is certainly related to the splittings used in algebraic multigrid (AMG [16]), where one seeks space decompositions depending on the operator at hand. In addition, both components of the splitting have a locally supported basis. As we will show, the analytical results and the preconditioning techniques obtained through the decomposition introduced in [9] for the Laplace equation can be extended to the case of problem (1.1) with jumps in the coefficients. With the orthogonal splitting of the DG space at hand, the solution of problem (1.1) reduces to solving two sub-problems; a non-conforming approximation to (1.1), and a problem in the complementary space containing *high oscillatory* error components. We show that the latter problem is easy to solve, since it is spectrally equivalent to its diagonal form, and therefore CG with a diagonal preconditioner is a uniform and robust solver. For the former problem, that is the approximation in the *low frequency* or Crouziex-Raviart space, we develop and analyze a two-level (overlapping) method and a BPX preconditioner. In fact, the theory for the BPX type multilevel preconditioner follows easily from the analysis of the two-level method.

We follow the approach taken in [78, 80] involving estimating the *asymptotic rate of convergence* of the preconditioned system. Nevertheless, dealing simultaneously with the jump in the coefficient κ and the non-nested character of the CR spaces presents extra difficulties in the analysis which preclude from a simple extension of the works [78, 80]. We are able to establish nearly optimal convergence and robustness (with respect to both the mesh size and the coefficient κ) for the two-level method and for the BPX preconditioner (up to a logarithmic term depending on the mesh size). The resulting algorithms involve the use of a solver in the CR space that is reduced to a smoothing step followed by conforming solver. Therefore, in particular one can argue that any of the robust and efficient solvers designed for conforming approximation of problem (1.1) could be used as a preconditioner here. The solvers are shown to be robust and uniformly convergent for symmetric IPDG-1 methods. For the non-symmetric IIPG and NIPG we propose two preconditioners; the symmetric part and a block-Jacobi-two level method. For the former, we indicate how the theory for second order problems in [9] can be successfully extended to cover problem (1.1), once the provisions given in this paper are taken into account. Unfortunately, for the latter, in spite of its simplicity, we are not able to provide a complete theory at this time, and it will be subject of future research.

Outline of the paper. The rest of the paper is organized as follows. We introduce the IPDG-1 and IPDG-0 methods for approximating (1.1) in §2 and revise some of their properties. The space decomposition of DG finite element space is introduced in §3. Consequences of the space splitting and solvers for the IPDG-0 methods are described in §4. The two-level and multi-level methods for the Crouziex-Raviart approximation are constructed and analyzed in Section5.2 (see §5.3). Section 6 is devoted to the design and analysis of the solvers for both the IPDG-0 and IPDG-1 family. Numerical experiments are included in §7, to verify the theory and assess the performance and robustness of the proposed preconditioners. The paper is completed with an Appendix where we have collected proofs of several technical results required in our analyses.

2. Discontinuous Galerkin Methods

In this section, we introduce the basic notation and describe the DG methods we consider for approximating problem (1.1). To begin, given a triangulation $\mathcal{T}_h$ of the domain Ω, we denote $\mathcal{E}_h := \mathcal{E}_h^o \cup \mathcal{E}_h^\partial$ the set of all edges (2D) /faces (3D) of $\mathcal{T}_h$, where $\mathcal{E}_h^o$ is the set of all interior edges/faces, and $\mathcal{E}_h^\partial$ is the set of all boundary edges/faces. Throughout the paper we shall use the standard notation for Sobolev spaces and their norms. We denote $H^2(\mathcal{T}_h)$ as the set of element-wise H^2 functions, and denote $L^2(\mathcal{E}_h)$ as the set of L^2 functions defined on $\mathcal{E}_h$. We will also use the notation $x_1 \lesssim y_1$, and $x_2 \gtrsim y_2$, whenever there exist constants C_1, C_2 independent

of the mesh size h and the coefficient κ or other parameters that x_1, x_2, y_1 and y_2 may depend on, and such that $x_1 \leq C_1 y_1$ and $x_2 \geq C_2 y_2$.

Trace Operators. Following [5], we recall the definition of the *average* and *jump* trace operators for scalar and vector-valued functions. Let T^+ and T^- be two neighboring elements, and $\mathbf{n}^+$, $\mathbf{n}^-$ be their outward normal unit vectors, respectively ($\mathbf{n}^\pm = \mathbf{n}_{T^\pm}$). Let $\zeta^\pm$ and $\boldsymbol{\tau}^\pm$ be the restriction of ζ and $\boldsymbol{\tau}$ to $T^\pm$. We set:

$$\begin{aligned} \{\!\{\zeta\}\!\} &= \frac{1}{2}(\zeta^+ + \zeta^-), \quad [\![\zeta]\!] = \zeta^+\mathbf{n}^+ + \zeta^-\mathbf{n}^- && \text{on } e \in \mathcal{E}_h^o, \\ \{\!\{\boldsymbol{\tau}\}\!\} &= \frac{1}{2}(\boldsymbol{\tau}^+ + \boldsymbol{\tau}^-), \quad [\![\boldsymbol{\tau}]\!] = \boldsymbol{\tau}^+ \cdot \mathbf{n}^+ + \boldsymbol{\tau}^- \cdot \mathbf{n}^- && \text{on } e \in \mathcal{E}_h^o, \end{aligned} \tag{2.1}$$

We also define the weighted average, $\{\!\{\cdot\}\!\}_\delta$, for any $\delta = \{\delta_e\}_{e\in\mathcal{E}_h^o}$, with $\delta_e \in [0,1]\ \forall\, e$, :

$$\{\!\{\zeta\}\!\}_\delta = \delta_e\zeta^+ + (1-\delta_e)\zeta^- \ , \quad \{\!\{\boldsymbol{\tau}\}\!\}_\delta = \delta_e\boldsymbol{\tau}^+ + (1-\delta_e)\boldsymbol{\tau}^- \ , \qquad \text{on } e \in \mathcal{E}_h^o \ . \tag{2.2}$$

For $e \in \mathcal{E}_h^\partial$, we set

$$[\![\zeta]\!] = \zeta\mathbf{n}, \quad \{\!\{\boldsymbol{\tau}\}\!\} = \{\!\{\boldsymbol{\tau}\}\!\}_\delta = \boldsymbol{\tau} \qquad \text{on } e \in \mathcal{E}_h^\partial. \tag{2.3}$$

We will also use the notation

$$(u,w)_{\mathcal{T}_h} = \sum_{T\in\mathcal{T}_h}\int_T uw dx \quad \forall\, u,w \in L^2(\Omega), \quad \langle u,w\rangle_{\mathcal{E}_h} = \sum_{e\in\mathcal{E}_h}\int_e uw \quad \forall\, u,w,\in L^2(\mathcal{E}_h).$$

The DG approximation to the model problem (1.1) can be written as

$$\text{Find } u_h \in V^{DG} \text{ such that } \mathcal{A}^{DG}(u,w) = (f,w)_{\mathcal{T}_h} \ , \quad \forall\, w \in V^{DG} \ , \tag{2.4}$$

where $\mathcal{A}^{DG}(\cdot,\cdot)$ is the bilinear form defining the method and can be of two different types. For the Type-1 family of weighted IP methods (see [68]) we set $\mathcal{A}^{DG}(\cdot,\cdot) = \mathcal{A}(\cdot,\cdot)$ where:

$$\begin{aligned} \mathcal{A}(u_h,w) = (\kappa\nabla u_h, \nabla w)_{\mathcal{T}_h} - \langle \{\!\{\kappa\nabla u_h\}\!\}_{\beta_e}, [\![w]\!]\rangle_{\mathcal{E}_h} + \theta\langle [\![u_h]\!], \{\!\{\kappa\nabla w\}\!\}_{\beta_e}\rangle_{\mathcal{E}_h} \\ + \langle \alpha_e h_e^{-1}\kappa_e [\![u_h]\!], [\![w]\!]\rangle_{\mathcal{E}_h}, \qquad \forall\, u_h,\, w \ \in V^{DG} \ . \end{aligned} \tag{2.5}$$

The weight $\beta = \{\beta_e\}_{e\in\mathcal{E}_h^o}$, depends on the coefficient κ and therefore it might vary over all interior edges/faces. For any $e \in \mathcal{E}_h^o$ with $e = \partial T^+ \cap \partial T^-$, we set β_e as follows:

$$\beta_e = \frac{\kappa^-}{\kappa^+ + \kappa^-} \quad \text{where} \quad \kappa^\pm = \kappa_{|_{T^\pm}}, \tag{2.6}$$

and we define the coefficient κ_e as the harmonic mean of κ^+ and κ^-:

$$\kappa_e := \frac{2\kappa^+\kappa^-}{\kappa^+ + \kappa^-} \ . \tag{2.7}$$

Remark 2.1. We remark that one could take κ_e as $\min\{\kappa^+,\kappa^-\}$ since both are equivalent:

$$\min\{\kappa^+,\kappa^-\} \leq \kappa_e = \frac{2\kappa^+\kappa^-}{\kappa^+ + \kappa^-} \leq 2\min\{\kappa^+,\kappa^-\} \leq \kappa^\pm \ . \tag{2.8}$$

The equivalence relations in (2.8) show that the results on spectral equivalence and uniform preconditioning given later on for (2.5) with κ_e defined in (2.7) (the harmonic mean), will automatically hold for method (2.5) with $\kappa_e := \min\{\kappa^+,\kappa^-\}$. To fix the notation and to simplify the presentation we use the harmonic mean as κ_e [see (2.7)].

The symmetric method was first considered in [68] and later in [37, Section 4] for variable coefficient (although there it was written using a slightly different notation and DG was only used in the skeleton of the partition). It was later extended to advection-diffusion problems in [25] and [32]. In (2.5), $\theta = -1$ gives the SIPG(β); $\theta = 1$ leads to NIPG(β); and $\theta = 0$ gives the IIPG(β) discretizations. The penalty parameter $\alpha_e > 0$ is set to a positive constant; and for

$\theta \neq 1$ should be taken large enough to ensure coercivity of the corresponding bilinear forms. We also introduce the corresponding family of IP(β)-0 methods, which use mid point quadrature rule for computing all the integrals in (2.5). That is, we set $\mathcal{A}^{DG}(\cdot,\cdot) = \mathcal{A}_0(\cdot,\cdot)$ with

$$\begin{aligned}\mathcal{A}_0(u_h,w) = (\kappa\nabla u_h,\nabla w)_{\mathcal{T}_h} - \langle \{\!\{\kappa\nabla u_h\}\!\}_{\beta_e},[\![w]\!]\rangle_{\mathcal{E}_h} + \theta\langle [\![u]\!],\{\!\{\kappa\nabla w\}\!\}_{\beta_e}\rangle_{\mathcal{E}_h} \\ + \langle \alpha_e h_e^{-1}\kappa_e \mathcal{P}_e^0([\![u_h]\!]),\mathcal{P}_e^0([\![w]\!])\rangle_{\mathcal{E}_h}, \qquad \forall\, u_h,\ w\ \in V^{DG}\,.\end{aligned} \tag{2.9}$$

where $\mathcal{P}_e^0 : L^2(\mathcal{E}_h) \mapsto \mathbb{P}^0(\mathcal{E}_h)$ is the L^2-projection onto the piecewise constants on $\mathcal{E}_h$. We note that this projection satisfies $\|\mathcal{P}_e^0\|_{L^2(\mathcal{E}_h)} = 1$.

Weighted Residual Formulation. Following [21] we can rewrite the two families of IP methods in the weighted residual framework: for all $u_h, w \in V^{DG}$,

$$\mathcal{A}(u_h,w) = (-\nabla\cdot(\kappa\nabla u_h),w)_{\mathcal{T}_h} + \langle [\![\kappa\nabla u_h]\!],\{\!\{w\}\!\}_{1-\beta_e}\rangle_{\mathcal{E}_h^o} + \langle [\![u_h]\!],\mathcal{B}_1(w)\rangle_{\mathcal{E}_h}, \tag{2.10}$$

$$\mathcal{A}_0(u_h,w) = (-\nabla\cdot(\kappa\nabla u_h),w)_{\mathcal{T}_h} + \langle [\![\kappa\nabla u_h]\!],\{\!\{w\}\!\}_{1-\beta_e}\rangle_{\mathcal{E}_h^o} + \langle [\![u_h]\!],\mathcal{P}_e^0(\mathcal{B}_1(w))\rangle_{\mathcal{E}_h}, \tag{2.11}$$

where $\mathcal{B}_1$ is defined as:

$$\mathcal{B}_1(w) = \theta\{\!\{\kappa\nabla w\}\!\}_{\beta_e} + \alpha_e h_e^{-1}\kappa_e [\![w]\!], \qquad \forall\, e\in\mathcal{E}_h. \tag{2.12}$$

Throughout the paper both the weighted residual formulation (2.10)-(2.11) and the standard one (2.5)-(2.9) will be used interchangeably.

Continuity and Coercivity. The family of methods (2.5) and (2.9) can be shown to provide an accurate and robust approximation to the solution of (1.1). We define the energy norm $|||\cdot|||_{DG}$:

$$|||u_h|||_{DG}^2 := \sum_{T\in\mathcal{T}_h}\kappa_T\|\nabla u_h\|_{0,T}^2 + \sum_{e\in\mathcal{E}_h}\kappa_e h_e^{-1}\|[\![u_h]\!]\|_{0,e}^2. \tag{2.13}$$

For the classical IP methods, the bilinear form $\mathcal{A}(\cdot,\cdot)$ is *continuous* and *coercive* in the above norm, with constants independent of the mesh size h and the coefficient κ:

$$\textit{Continuity:} \qquad |\mathcal{A}(u_h,w)| \lesssim |||u_h|||_{DG}\,|||w|||_{DG}\,, \qquad \forall\, u_h\,,\ w\ \in V^{DG} \tag{2.14}$$

$$\textit{Coercivity:} \qquad \mathcal{A}(u_h,u_h) \gtrsim |||u_h|||_{DG}^2\,, \qquad \forall u_h \in V^{DG}\,. \tag{2.15}$$

The proof of (2.15) and (2.14) is standard and could be found in [37]. We sketch it here for completeness. Note first that for each $e\in\mathcal{E}_h^o$, such that $e = \partial T^+\cap\partial T^-$, the weighted average $\{\!\{\kappa\nabla u\}\!\}_{\beta_e}$ can be rewritten as:

$$\begin{aligned}\{\!\{\kappa\nabla u\}\!\}_{\beta_e} &= \beta_e(\kappa^+(\nabla u)^+) + (1-\beta_e)(\kappa^-(\nabla u)^-) \\ &= \frac{\kappa^-}{\kappa^+ + \kappa^-}\kappa^+(\nabla u)^+ + \frac{\kappa^+}{\kappa^+ + \kappa^-}\kappa^-(\nabla u)^- \\ &= \frac{\kappa^+\kappa^-}{\kappa^+ + \kappa^-}[(\nabla u)^+ + (\nabla u)^-] = \kappa_e\{\!\{\nabla u\}\!\}\,.\end{aligned} \tag{2.16}$$

The trace inequality [1], the inverse inequality [28], and (2.8) then imply the following bounds

$$\begin{aligned}h_e\|\{\!\{\kappa\nabla u\}\!\}_{\beta_e}\|_{0,e}^2 &\le C_t(\kappa_e)^2\left(\|\nabla u\|_{0,T^+\cup T^-}^2 + h^2|\nabla u|_{1,T^+\cup T^-}^2\right) \\ &\le 2(\kappa_e)C_t(1+C_{inv}^2)\left(\kappa^+\|\nabla u\|_{0,T^+}^2 + \kappa^-\|\nabla u\|_{0,T^-}^2\right)\end{aligned} \tag{2.17}$$

Then, using Cauchy-Schwarz inequality, the estimate (2.17), and the arithmetic-geometric inequality we have

$$\begin{aligned}|\langle\{\!\{\kappa\nabla u\}\!\}_{\beta_e},[\![w]\!]\rangle_{\mathcal{E}_h}| &= \left|\sum_{e\in\mathcal{E}_h}\int_e \kappa_e\{\!\{\nabla u\}\!\}[\![w]\!]ds\right| \\ &\le \left(\sum_{e\in\mathcal{E}_h}\frac{1}{\alpha_e}h_e\kappa_e\|\{\!\{\nabla u\}\!\}\|^2_{0,e}\right)^{1/2}\left(\sum_{e\in\mathcal{E}_h}\alpha_e h_e^{-1}\kappa_e\|[\![w]\!]\|^2_{0,e}\right) \\ &\le \frac{8C_t(1+C_{inv}^2)}{\alpha_e}\sum_{T\in\mathcal{T}_h}\kappa_T\|\nabla u\|^2_{0,T}+\frac{1}{4}\sum_{e\in\mathcal{E}_h}\alpha_e h_e^{-1}\kappa_e\|[\![w]\!]\|^2_{0,e}\end{aligned}$$

Then, (2.14) follows from Cauchy-Schwarz inequality and the estimate (2.18). The inequality (2.15) is proved by setting $w=u$ in (2.5) and taking into account the estimate (2.18) (also with $w=u$). We have,

$$\begin{aligned}\mathcal{A}(u,u) &= \sum_{T\in\mathcal{T}_h}\kappa_T\|\nabla u\|^2_{0,T}+\alpha_e\sum_{e\in\mathcal{E}_h}\kappa_e h_e^{-1}\|[\![u]\!]\|^2_{0,e}-(1-\theta)\langle\{\!\{\kappa\nabla u\}\!\}_{\beta_e},[\![u]\!]\rangle_{\mathcal{E}_h} \\ &\ge |||u|||^2_{DG}-|1-\theta|\,|\langle\{\!\{\kappa\nabla u\}\!\}_{\beta_e},[\![u]\!]\rangle_{\mathcal{E}_h}| \\ &\ge \left(1-\frac{8C_t(1+C_{inv}^2)}{\alpha_e}\right)\sum_{T\in\mathcal{T}_h}\kappa_T\|\nabla u\|^2_{0,T}+\frac{4-|1-\theta|}{4}\alpha_e\sum_{e\in\mathcal{E}_h}\kappa_e h_e^{-1}\|[\![u]\!]\|^2_{0,e}\,,\end{aligned}$$

and (2.15) follows immediately by taking α_e large enough (if $\theta\neq 1$). Moreover, notice that both constants in (2.14) and (2.15) depend on the shape regularity of the mesh partition but are independent of the coefficient κ.

For the IP-0 methods (2.9), similar properties can also be easily shown to hold following the same arguments, albeit in a different energy norm, defined as:

$$|||u|||^2_{DG0} := \sum_{T\in\mathcal{T}_h}\kappa_T\|\nabla u\|^2_{0,T}+\sum_{e\in\mathcal{E}_h}\kappa_e h_e^{-1}\|\mathcal{P}^0_e([\![u]\!])\|^2_{0,e}. \tag{2.18}$$

For both families of methods, optimal error estimates in the energy norms (2.13) and (2.18) can be shown, arguing as in [5]. See also [8] for further discussion on the L^2-error analysis of these methods.

We now establish the spectral equivalence between $\mathcal{A}(\cdot,\cdot)$ and $\mathcal{A}_0(\cdot,\cdot)$.

Lemma 2.2. *Let $\mathcal{A}(\cdot,\cdot)$ be a bilinear form corresponding to a Type-1 IP(β) method* (2.5) *and let $\mathcal{A}_0(\cdot,\cdot)$ be the corresponding Type-0 bilinear form as defined in* (2.9)*. Then there exists a positive constant $c_0=c_0(\alpha)$, depending only on the shape regularity of the mesh and the penalty parameter α (but independent of the coefficient κ and the mesh size h) such that,*

$$\mathcal{A}_0(u,u)\le\mathcal{A}(u,u)\le c_0(\alpha)\mathcal{A}_0(u,u)\quad\forall u\in V^{DG}. \tag{2.19}$$

Proof. The lower bound follows immediately from the fact that the projection $\mathcal{P}^0_e$ is an $L^2(e)$-orthogonal projection and therefore has norm 1.

To show the upper bound it is enough to show that

$$\sum_{e\in\mathcal{E}_h}\alpha_e h_e^{-1}\kappa_e\|[\![u]\!]\|^2_{0,e}\le C(\sum_{T\in\mathcal{T}_h}\kappa_T\|\nabla u\|^2_{0,T}+\sum_{e\in\mathcal{E}_h}\alpha_e h_e^{-1}\kappa_e\|\mathcal{P}^0_e[\![u]\!]\|^2_{0,e}). \tag{2.20}$$

Adding and subtracting $\mathcal{P}^0_e[\![u]\!]$ in the term on the left side above and using again that $\mathcal{P}^0_e$ is the $L_2(e)$-orthogonal projection on the constant functions we have for each face e,

$$\|[\![u]\!]\|^2_{0,e}=\|\mathcal{P}^0_e([\![u]\!])\|^2_{0,e}+\|[\![u]\!]-\mathcal{P}^0_e([\![u]\!])\|^2_{0,e}\,. \tag{2.21}$$

Hence, we only need to estimate the last term on the right side of the above inequality and show that this term is bounded by the right side in the inequality given in (2.20). Observe that on each $e \in \mathcal{E}$, $e = \partial T^+ \cap \partial T^-$,

$$\|[\![u]\!] - \mathcal{P}_e^0([\![u]\!])\|_{0,e}^2 \lesssim \|u^+ - \mathcal{P}_e^0(u^+)\|_{0,e}^2 + \|u^- - \mathcal{P}_e^0(u^-)\|_{0,e}^2 .$$

In [9, Lemma 2.1] it was further shown that

$$h_e^{-1}\|u^\pm - \mathcal{P}_e^0(u^\pm)\|_{0,e}^2 \lesssim |u^\pm|_{1,T^\pm}^2 . \quad \forall\, e \in \mathcal{E}_h$$

Thus, multiplying the above expression by the harmonic average κ_e and taking into account (2.8) we arrive at

$$\sum_{e\in\mathcal{E}_h} \alpha_e h_e^{-1}\kappa_e \|[\![u]\!]\|_{0,e}^2 \le C(\alpha) \sum_{T\in\mathcal{T}_h} \kappa_T \|\nabla u\|_{0,T}^2 + \sum_{e\in\mathcal{E}_h} \alpha_e h_e^{-1}\kappa_e \|\mathcal{P}_e^0[\![u]\!]\|_{0,e}^2 ,$$

which concludes the proof. □

3. Space decomposition of the V^{DG} space

In this section we introduce a decomposition of the V^{DG}-space that will play a key role in the design of the solvers for the DG discretizations (2.5) and (2.9). In [9] (see also [24]) it is shown that the discontinuous piecewise linear finite element space V^{DG} admits the decomposition: $V^{DG} = V_h^{CR} \oplus \mathcal{Z}$ where V_h^{CR} denotes the standard Crouziex-Raviart space defined as

$$V_h^{CR} = \left\{ v \in L^2(\Omega) \,:\, v_{|_T} \in \mathbb{P}^1(T)\ \forall T \in \mathcal{T}_h \text{ and } \mathcal{P}_e^0([\![v]\!]\cdot \mathbf{n}) = 0\ \forall\, e \in \mathcal{E}_h^o \right\}, \tag{3.1}$$

and the complementary space $\mathcal{Z}$ is a space containing functions having zero averages at the midpoints of the internal edges:

$$\mathcal{Z} = \left\{ z \in L^2(\Omega) \,:\, z_{|_T} \in \mathbb{P}^1(T)\ \forall T \in \mathcal{T}_h \text{ and } \mathcal{P}_e^0\{\!\{v\}\!\} = 0,\ \forall\, e \in \mathcal{E}_h^o \right\}.$$

In [9] it was shown that this decomposition satisfies $\mathcal{A}_0(v, z) = 0$, for all $v \in V_h^{CR}$ and $z \in \mathcal{Z}$. We now modify the definition of $\mathcal{Z}$ above in order to account for the presence of a coefficient in the problem (1.1); we define

$$\mathcal{Z}_\beta = \left\{ z \in L^2(\Omega) \,:\, z_{|_T} \in \mathbb{P}^1(T)\ \forall T \in \mathcal{T}_h \text{ and } \mathcal{P}_e^0(\{\!\{z\}\!\}_{1-\beta_e}) = 0,\ \forall\, e \in \mathcal{E}_h^o \right\}, \tag{3.2}$$

where the weight β was defined earlier in (2.6). Note that the weight β depends on the coefficient κ, and, as a consequence, the space $\mathcal{Z}_\beta$ is also coefficient dependent. In what follows, we shall show that $\mathcal{Z}_\beta$ is a space complementary to V_h^{CR} in V^{DG} and the corresponding decomposition has properties analogous to the properties of the decomposition $V^{DG} = V_h^{CR} \oplus \mathcal{Z}$ given in [9] for the Poisson problem.

For any $e \in \mathcal{E}_h$ with $e \subset T \in \mathcal{T}_h$, let $\varphi_{e,T}$ be the canonical Crouzeix-Raviart basis function on T, which is defined by

$$\varphi_{e,T}|_T \in \mathbb{P}^1(T), \quad \varphi_{e,T}(m_{e'}) = \delta_{e,e'}\ \ \forall e' \in \mathcal{E}_h(T), \text{ and } \varphi_{e,T}(x) = 0\ \ \forall x \notin T,$$

where m_e is the mass center e. We will denote n_T and n_E as the number of simplices and faces (or edges when $d = 2$) respectively. We also denote n_{BE} as the number of boundary faces.

Proposition 3.1. *For any $u \in V^{DG}$ there exists a unique $v \in V_h^{CR}$ and a unique $z_{\beta_e} \in \mathcal{Z}_\beta$ such that $u = v + z_{\beta_e}$, that is*

$$V^{DG} = V_h^{CR} \oplus \mathcal{Z}_\beta. \tag{3.3}$$

Proof. Throughout the proof, for simplicity, let us set $\beta^+ = \beta_e$, $\beta^- = (1 - \beta_e)$, and $\varphi_e^\pm = \varphi_{e,T^\pm}$ for any $e \in \mathcal{E}_h^o$ with $e = \partial T^+ \cap \partial T^-$. We also denote $\varphi_e = \varphi_{e,T}$ for any $e \in \mathcal{E}_h^\partial$ with $e = \partial T \cap \partial\Omega$. Note that

$$\dim V^{DG} = (d+1)n_T = 2n_E - n_{BE},$$

and it is also obvious that $\{\varphi_e^{\pm}\}_{e\in\mathcal{E}_h^o}\cup\{\varphi_e\}_{e\in\mathcal{E}_h^\partial}$ form a basis for V^{DG}. Notice that $\beta^+ + \beta^- = 1$, therefore we can express any $u \in V^{DG}$ as

$$\begin{aligned}
u(x) &= \sum_{e\in\mathcal{E}_h^o} u^+(m_e)\varphi_e^+(x) + \sum_{e\in\mathcal{E}_h^o} u^-(m_e)\varphi_e^-(x) + \sum_{e\in\mathcal{E}_h^\partial} u(m_e)\varphi_e(x)\\
&= \sum_{e\in\mathcal{E}_h^o} (\beta^- u^+(m_e) + \beta^+ u^-(m_e))(\varphi_e^+(x) + \varphi_e^-(x))\\
&\quad + \sum_{e\in\mathcal{E}_h^o} (u^+(m_e) - u^-(m_e))(\beta^+\varphi_e^+(x) - \beta^-\varphi_e^-(x)) + \sum_{e\in\mathcal{E}_h^\partial} u(m_e)\varphi_e(x)\\
&= \sum_{e\in\mathcal{E}_h^o} \left(\frac{1}{|e|}\int_e \{\!\{u\}\!\}_{1-\beta_e} ds\right)(\varphi_e^+(x) + \varphi_e^-(x))\\
&\quad + \sum_{e\in\mathcal{E}_h^o} \left(\frac{1}{|e|}\int_e [\![u]\!]\mathbf{n}^+ ds\right)(\beta^+\varphi_e^+(x) - \beta^-\varphi_e^-(x)) + \sum_{e\in\mathcal{E}_h^\partial} \left(\frac{1}{|e|}\int_e [\![u]\!]\mathbf{n} ds\right)\varphi_e(x)\\
&= v(x) + z_\beta(x).
\end{aligned}$$

Then for each $e \in \mathcal{E}_h^o$, we set

$$\varphi_e^{CR}(x) := \varphi_e^+(x) + \varphi_e^-(x), \tag{3.4}$$

and let

$$\psi_e^z(x) := \beta^+\varphi_e^+(x) - \beta^-\varphi_e^-(x) = \begin{cases} \beta^+\varphi_e^+(x), & x \in T^+ \\ -\beta^-\varphi_e^-(x), & x \in T^- \end{cases}, \tag{3.5}$$

and $\psi_e^z(x) := 0$ for all $x \notin T^+ \cup T^-$. In the definition (3.5) of $\psi_e^z(x)$, we have used $\varphi_e^-(x) = 0$ for $x \in T^+$ and $\varphi_e^+(x) = 0$ for $x \in T^-$. Finally, when $e \in \mathcal{E}_h^\partial$ with $e = \partial T \cap \partial\Omega$ for some T, we set

$$\psi_e^z(x) = \varphi_e(x), \quad \forall x \in T. \tag{3.6}$$

It is then straightforward to check that

$$V_h^{CR} = \operatorname{span}\{\varphi_e^{CR}\}_{e\in\mathcal{E}_h^o}, \quad \text{and} \quad \mathcal{Z}_\beta = \operatorname{span}\{\psi_e^z\}_{e\in\mathcal{E}_h}.$$

Hence, for all $u \in V^{DG}$ there are unique $v \in V_h^{CR}$ and $z_{\beta_e} \in \mathcal{Z}_\beta$ defined by

$$\begin{aligned}
v &= \sum_{e\in\mathcal{E}_h^o} \left(\frac{1}{|e|}\int_e \{\!\{u\}\!\}_{1-\beta_e} ds\right)\varphi_e^{CR}(x) \in V_h^{CR},\\
z_\beta &= \sum_{e\in\mathcal{E}_h} \left(\frac{1}{|e|}\int_e [\![u]\!]\mathbf{n}^+ ds\right)\psi_e^z(x) \in \mathcal{Z}_\beta.
\end{aligned}$$

such that $u = v + z_\beta$. This shows (3.3)and concludes the proof. □

Remark 3.2. As we pointed out in the introduction, the definition of the subspace $\mathcal{Z}_\beta$ clearly depends on the coefficient κ since β depends κ. Such dependence is also often seen in algebraic multigrid analysis, where the coarse spaces depend on the operator, and are in fact constructed in this way, the aim being to increase robustness of the methods.

In the proof of Proposition 3.1 above we have introduced the basis in both V_h^{CR} and $\mathcal{Z}_\beta$. The canonical Crouzeix-Raviart basis functions are denoted with $\{\varphi_e^{CR}\}_{e\in\mathcal{E}_h^o}$ for non-conforming and these functions are continuous at the mass centers m_e of the faces $e \in \mathcal{E}_h^o$. The basis in $\mathcal{Z}_\beta$, $\{\psi_e^z\}_{e\in\mathcal{E}_h}$ consists of piecewise $\mathbb{P}^1$ functions, which are discontinuous across the faces in $\mathcal{E}_h$. In fact, for any $z \in \mathcal{Z}_\beta$ such that $z = \sum_{e\in\mathcal{E}_h} z_e\psi_e^z$ with $z_e \in \mathbb{R}$ for any $e \in \mathcal{E}_h$, we have

$$([\![z]\!]\mathbf{n}^+)(m_{e'}) = z_{e'}, \qquad \forall e' \in \mathcal{E}_h.$$

To see this, evaluating the jump of z at $m_{e'}$ gives

$$\begin{aligned}(\llbracket z\rrbracket\mathbf{n}^+)(m_{e'}) &= \sum_{e\in\mathcal{E}_h} z_e(\llbracket\psi_e^z\rrbracket\mathbf{n}^+)(m_{e'}) = z_{e'}(\llbracket\psi_{e'}^z\rrbracket\mathbf{n}^+)(m_{e'})\\ &= \begin{cases} z_{e'}(\beta_{e'}-(\beta_{e'}-1)) = z_{e'}, & e'\in\mathcal{E}_h^o,\\ z_{e'}, & e'\in\mathcal{E}_h^\partial. \end{cases}\end{aligned}$$

This relation will also be used later to obtain uniform diagonal preconditioners for the restrictions of $\mathcal{A}(\cdot,\cdot)$ and $\mathcal{A}_0(\cdot,\cdot)$ on $\mathcal{Z}_\beta$.

Remark 3.3. For mixed boundary value problems, that is, $\partial\Omega$ contains both Neumann boundary $\Gamma_N\neq\emptyset$ and Dirichlet boundary Γ_D with $\partial\Omega=\Gamma_D\cup\Gamma_N$, the definition of the basis functions on the boundary faces [see (3.6)] need to be changed as:

$$\begin{array}{rcll}\psi_e^{CR}(x) &=& \varphi_{e,T}(x), & e=\partial T\cap\Gamma_N, \quad \text{for all} \quad x\in T,\\ \psi_e^{z}(x) &=& \varphi_{e,T}(x), & e=\partial T\cap\Gamma_D, \quad \text{for all} \quad x\in T.\end{array} \tag{3.7}$$

Thus, in case $\Gamma_N\neq\emptyset$ the dimension of V_h^{CR} is increased (by adding to it functions that correspond to degrees of freedom on Γ_N) and the dimension of $\mathcal{Z}_\beta$ is decreased accordingly. Clearly things balance out correctly: the identity $V^{DG}=V_h^{CR}\oplus\mathcal{Z}_\beta$ holds, and also the analysis carries out with very little modification.

Next lemma is a simple but a key observation used in the design of efficient solvers and preconditioners and it shows that the restriction of $\mathcal{A}_0(\cdot,\cdot):V^{DG}\times V^{DG}\mapsto\mathbb{R}$ to $V_h^{CR}\times\mathcal{Z}_\beta$ vanishes.

Lemma 3.4. *Let $u\in V^{DG}$ be such that $u=v+z$ with $v\in V_h^{CR}$ and $z\in\mathcal{Z}_\beta$. Let $\mathcal{A}_0(\cdot,\cdot)$ be the bilinear form defined in* (2.9)*. Then,*

$$\mathcal{A}_0(v,z)=0 \qquad \forall\, v\in V_h^{CR}, \quad \forall\, z\in\mathcal{Z}_\beta. \tag{3.8}$$

Furthermore if $\mathcal{A}_0(\cdot,\cdot)$ is symmetric, then $\mathcal{A}_0(v,z)=\mathcal{A}_0(z,v)=0$ for all $v\in V_h^{CR}$, and for all $z\in\mathcal{Z}_\beta$ and the decomposition (3.3) *is $\mathcal{A}_0$-orthogonal, namely, $V_h^{CR}\ \perp_{\mathcal{A}_0}\ \mathcal{Z}_\beta$.*

Proof. From the weighted-residual form of $\mathcal{A}_0(\cdot,\cdot)$ given in (2.11), for all $v\in V_h^{CR}$, and all $z\in\mathcal{Z}_\beta$ we easily obtain

$$\mathcal{A}_0(v,z)=(-\nabla\cdot(\kappa\nabla v),z)_{\mathcal{T}_h}+\langle\llbracket\kappa\nabla v\rrbracket,\{\!\{z\}\!\}_{1-\beta_e}\rangle_{\mathcal{E}_h^o}+\langle\llbracket v\rrbracket,\mathcal{P}_e^0(\mathcal{B}_1(z))\rangle_{\mathcal{E}_h}=0.$$

In the equation above, the first term is zero due to the fact that v is linear in each T and that the coefficient κ is a constant on T; the second term vanishes from the definition of $\mathcal{Z}_\beta$ (since $\llbracket\kappa\nabla v\rrbracket$ is constant on each $e\in\mathcal{E}_h^o$) and last term vanishes as well independently of the choice of θ (or equivalently the choice of $\mathcal{B}_1(v)$) from the definition of the space V_h^{CR}. Moreover, in the case when $\mathcal{A}_0(\cdot,\cdot)$ is symmetric and positive definite we have that $\mathcal{A}_0(v,z)=\mathcal{A}_0(z,v)$, for all $v\in V_h^{CR}$ and for all $z\in\mathcal{Z}_\beta$. Thus, for the symmetric method $\mathcal{A}_0(\cdot,\cdot)$, the spaces V_h^{CR} and $\mathcal{Z}_\beta$ are indeed $\mathcal{A}_0$-orthogonal. The proof is complete. □

4. Solvers for IP-0 methods

In this section we show how Proposition 3.1 and Lemma 3.4 can be used in the design and analysis of uniformly convergent iterative methods for the IP-0 methods. We follow the ideas and analysis introduced in [9] and point out the differences. We first consider the approximation to problem (1.1) with $\mathcal{A}^{DG}(\cdot,\cdot)=\mathcal{A}_0(\cdot,\cdot)$. To begin, let A_0 be the discrete operator defined by $(A_0u,w)=\mathcal{A}_0(u,w)$ and let $\mathbb{A}_0$ be its matrix representation in the new basis (3.7). We denote by $\mathbf{u}=[\mathbf{z},\mathbf{v}]^T$, $\mathbf{f}=[\mathbf{f}_\mathbf{z},\mathbf{f}_\mathbf{v}]^T$ be the vector representation of the unknown function u and of the

right hand side f, respectively, in this new basis. A simple consequence of Lemma 3.4 is that the matrix $\mathbb{A}_0$ (in this basis) has block lower triangular structure:

$$\mathbb{A}_0 = \begin{bmatrix} \mathbb{A}_0^{zz} & \mathbf{0} \\ \mathbb{A}_0^{vz} & \mathbb{A}_0^{vv} \end{bmatrix}. \tag{4.1}$$

where $\mathbb{A}_0^{zz}, \mathbb{A}_0^{vv}$ are the matrix representation of A_0 restricted to the subspaces $\mathcal{Z}_\beta$ and V^{CR}, respectively, and $\mathbb{A}_0^{vz}$ is the matrix representation of the term that accounts for the coupling (or non-symmetry) $\mathcal{A}_0(\psi^z, \varphi^{CR})$. As remarked earlier, for Type-0 SIPG(β) discretization, the stiffness matrix $\mathbb{A}_0$ is block-diagonal.

Here, we have given a 2D example, with two squares $\Omega_1 = [-0.5, 0]^2$ and $\Omega_2 = [0, 0.5]^2$ inside the domain $\Omega = [-1, 1]^2$. We set the coefficients $\kappa(x) = 1$ for all $x \in \Omega_1 \cup \Omega_2$ and $\kappa(x) = 10^{-3}$ when $x \in \Omega \setminus (\Omega_1 \cup \Omega_2)$, see Figure 4.1. We choose the penalty parameter $\alpha_e = 20$ in (2.9). Figure 4.2 shows the sparsity patterns of the Type-0 IP(β) methods with standard nodal basis,

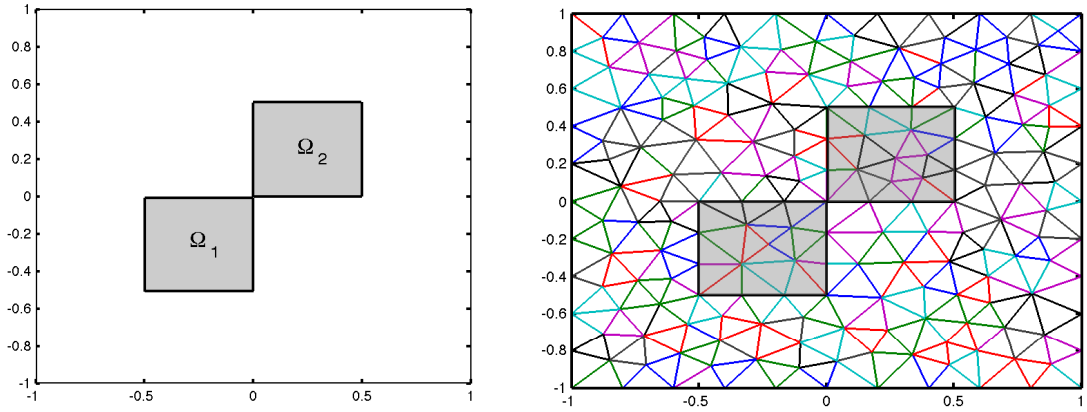


FIGURE 4.1. *Computational domain and unstructured mesh.*

while Figure 4.3 shows the sparsity patterns of the stiffness matrices for the IP-0 (β) methods after changing from the standard nodal basis to the basis (3.4)-(3.5) induced by the splitting $V^{DG} = V_h^{CR} \oplus \mathcal{Z}_\beta$.

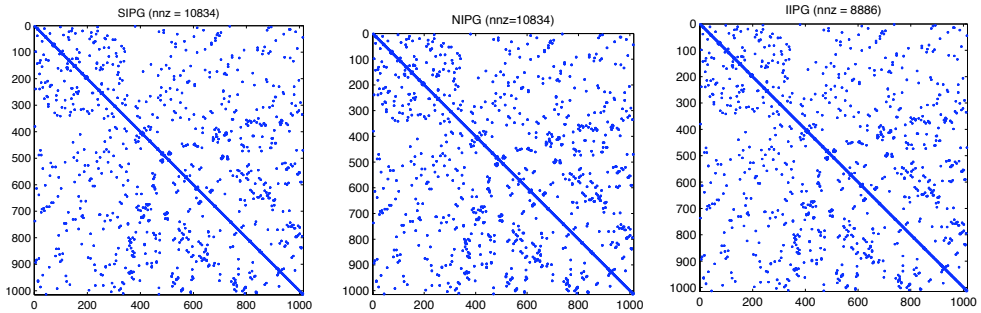


FIGURE 4.2. Non-zero pattern of the matrix representation in the standard nodal basis of the operators associated with Type-0 IP methods. From left to right: SIPG, NIPG and IIPG methods.

Clearly, as in the constant coefficient case, a simple algorithm based on a block version of forward substitution provides an exact solver for the solution of the linear systems with coefficient matrix $\mathbb{A}_0$. A formal description of this block forward substitution is given in the next Algorithm.

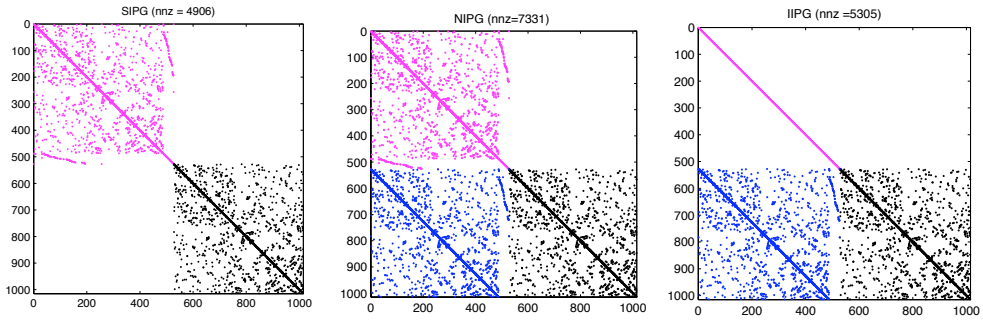


FIGURE 4.3. Non-zero pattern of the matrix representations according to the basis (3.3) of the operator associated with Type-0 IP methods;i.e., $\mathbb{A}_0$. From left to right: SIPG, NIPG and IIPG methods.

Algorithm 4.1.
1. Find $z \in \mathcal{Z}_\beta$ such that $\mathcal{A}_0(z,\psi) = (f,\psi)_{\mathcal{T}_h}$ for all $\psi \in \mathcal{Z}_\beta$
2. Find $v \in V^{CR}$ such that $\mathcal{A}_0(v,\varphi) = (f,\varphi)_{\mathcal{T}_h} - \mathcal{A}_0(z,\varphi)$ for all $\varphi \in V^{CR}$
3. Set $u = z + v$

Notice that the above algorithm requires the solution of $\mathcal{A}_0(\cdot,\cdot)$ on $\mathcal{Z}_\beta$ (Step 1. of the algorithm) and the solution of $\mathcal{A}_0(\cdot,\cdot)$ on V_h^{CR} (Step 2. of Algorithm 4.1). Unlike the situation in [9], due to the jump coefficient in (1.1), the solution on V^{CR} is more involved, and therefore we postpone its discussion and analysis until Section 5. On the other hand, as we show in the next section, the solution on $\mathcal{Z}_\beta$ (the space regarded as containing "high frequency" components) can be handled efficiently using CG with a diagonal preconditioner.

4.1. **Solution on $\mathcal{Z}_\beta$.** The first result in this section establishes the symmetry of the restrictions of the bilinear forms (of both Type-0 and Type-1) on $\mathcal{Z}_\beta$.

Lemma 4.2. *Let $\mathcal{A}(\cdot,\cdot)$ be the bilinear form of a non-symmetric Type-1 IP method as defined in (2.5) and let $\mathcal{A}_0(\cdot,\cdot)$ be the corresponding Type-0 bilinear form. Then the restrictions to $\mathcal{Z}_\beta$ of both $\mathcal{A}_0(\cdot,\cdot)$ and $\mathcal{A}(\cdot,\cdot)$ are symmetric. Namely, for $\theta = -1, 0, 1$ and for all $z \in \mathcal{Z}_\beta$ and $\phi \in \mathcal{Z}_\beta$ we have*

$$\mathcal{A}_0(z,\phi) = \mathcal{A}_0(\phi,z) \quad \text{and} \quad \mathcal{A}(z,\phi) = \mathcal{A}(\phi,z).$$

Proof. If $\theta = -1$ there is nothing to prove, since in this case both bilinear forms are symmetric. Hence we only consider the cases $\theta = 0$ or $\theta = -1$. Integrating by parts and using the fact that $z \in \mathcal{Z}_\beta$ and $\psi \in \mathcal{Z}_\beta$ are linear on each element T shows that

$$0 = (-\nabla\cdot(\kappa\nabla\psi), \nabla z)_{\mathcal{T}_h} = (\kappa\nabla\psi,\nabla z)_{\mathcal{T}_h} - \langle \{\!\{\kappa\nabla\psi\}\!\}_{\beta_e}, [\![z]\!]\rangle_{\mathcal{E}_h} - \langle [\![\kappa\nabla\psi]\!], \{\!\{z\}\!\}_{1-\beta_e}\rangle_{\mathcal{E}_h^o}\,. \tag{4.2}$$

Hence, from the definition (3.2) of the $\mathcal{Z}_\beta$ space, it follows that

$$(\kappa\nabla\psi,\nabla z)_{\mathcal{T}_h} = \langle \{\!\{\kappa\nabla\psi\}\!\}_{\beta_e}, [\![z]\!]\rangle_{\mathcal{E}_h} = \langle \{\!\{\kappa\nabla z\}\!\}_{\beta_e}, [\![\psi]\!]\rangle_{\mathcal{E}_h} \quad \forall\, z,\psi \in \mathcal{Z}_\beta. \tag{4.3}$$

Substituting the above identity in the definition of the bilinear form (2.9) then leads to

$$\begin{aligned}\mathcal{A}_0(z,\psi) &= \theta\langle [\![z]\!], \{\!\{\kappa\nabla\psi\}\!\}_{\beta_e}\rangle_{\mathcal{E}_h} + \langle \mathcal{P}_e^0([\![z]\!]), \kappa_e[\![\psi]\!]\rangle_{\mathcal{E}_h} \\ &= \theta(\kappa\nabla\psi,\nabla z)_{\mathcal{T}_h} + \langle \mathcal{P}_e^0([\![\psi]\!]), \kappa_e[\![z]\!]\rangle_{\mathcal{E}_h} = \mathcal{A}_0(\psi,z).\end{aligned}$$

This shows the symmetry of $\mathcal{A}_0(\cdot,\cdot)$ on $\mathcal{Z}_\beta$. The symmetry for $\mathcal{A}(\cdot,\cdot)$ on $\mathcal{Z}_\beta$ then follows from the result for $\mathcal{A}_0(\cdot,\cdot)$, since the difference between these bilinear forms is obviously symmetric. □

We now study the conditioning of the bilinear forms $\mathcal{A}(\cdot,\cdot)$ and $\mathcal{A}_0(\cdot,\cdot)$ on $\mathcal{Z}_\beta$. For all $z \in \mathcal{Z}_\beta$, and for all $\phi \in \mathcal{Z}_\beta$ with

$$z = \sum_{e\in\mathcal{E}_h} z_e\psi_e^z \ \in \ \mathcal{Z}_\beta, \quad \text{and} \quad \phi = \sum_{e\in\mathcal{E}_h} \phi_e\psi_e^z \ \in \ \mathcal{Z}_\beta.$$

we introduce a weighted scalar product $(\cdot,\cdot)_* : \mathcal{Z}_\beta \times \mathcal{Z}_\beta \mapsto \mathbb{R}$ and the corresponding norm $\|\cdot\|_*$, defined as follows

$$(z,\phi)_* := \sum_{e\in\mathcal{E}_h} \frac{|e|}{h_e}\kappa_e z_e\phi_e \ , \quad \|z\|_*^2 := (z,z)_*. \tag{4.4}$$

Observe that the matrix representation of the above weighted scalar product (in the basis given in (3.5)), is in fact a diagonal matrix. The next result shows that the restrictions of $\mathcal{A}(\cdot,\cdot)$ and $\mathcal{A}_0(\cdot,\cdot)$ to $\mathcal{Z}_\beta$ are spectrally equivalent to the weighted scalar product $(\cdot,\cdot)_*$ and therefore their matrix representations are spectrally equivalent to a diagonal matrix.

Lemma 4.3. *Let $\mathcal{Z}_\beta$ be the space defined in* (3.2)*. Then for all $z \in \mathcal{Z}_\beta$, the following estimates hold*

$$\|z\|_*^2 \lesssim \mathcal{A}_0(z,z) \lesssim \|z\|_*^2, \tag{4.5}$$

and also

$$\|z\|_*^2 \lesssim \mathcal{A}(z,z) \lesssim \|z\|_*^2. \tag{4.6}$$

Proof. Let us fix $z \in \mathcal{Z}_\beta$, $z = \sum_{e\in\mathcal{E}_h} z_e\psi_e^z$. From the definition of $\mathcal{P}_e^0([\![z]\!])\|_{0,e}^2$, it is immediate to see that

$$\|\mathcal{P}_e^0([\![z]\!])\|_{0,e}^2 = |e|z_e^2.$$

Thus, we have that

$$\sum_{e\in\mathcal{E}_h} \kappa_e h_e^{-1}\|\mathcal{P}_e^0([\![z]\!])\|_{0,e}^2 = \sum_{e\in\mathcal{E}_h} \kappa_e\frac{|e|}{h_e}z_e^2. \tag{4.7}$$

From this relation it is easy to show both estimates (4.5) and (4.6). First, to show (4.5), we notice that taking into account (4.3) together with the estimate (2.8), it follows that

$$\sum_{T\in\mathcal{T}_h} \kappa_T\|\nabla z\|_{0,T}^2 = (\kappa\nabla z,\nabla z)_{\mathcal{T}_h} = \langle \{\!\{\kappa\nabla z\}\!\}_{\beta_e}, [\![z]\!]\rangle_{\mathcal{E}_h} = \langle \kappa_e\{\!\{\nabla z\}\!\}, \mathcal{P}_e^0([\![z]\!])\rangle_{\mathcal{E}_h}$$

$$\lesssim \left(\sum_{T\in\mathcal{T}_h} \kappa_T\|\nabla z\|_{0,T}^2\right)^{1/2} \left(\sum_{e\in\mathcal{E}_h} \kappa_e\|h_e^{-1/2}\mathcal{P}_e^0([\![z]\!])\|_{0,e}^2\right)^{1/2}$$

and therefore,

$$\sum_{T\in\mathcal{T}_h} \kappa_T\|\nabla z\|_{0,T}^2 \lesssim \sum_{e\in\mathcal{E}_h} \kappa_e\|h_e^{-1/2}\mathcal{P}_e^0([\![z]\!])\|_{0,e}^2 = \|z\|_*^2, \tag{4.8}$$

and since $z \in \mathcal{Z}_\beta$ was arbitrary, we have that $\mathcal{A}_0(z,z) \lesssim \|z\|_*^2$, for all $z \in \mathcal{Z}_\beta$ and this proves the upper bound in (4.5).

To prove the lower bound, we use the coercivity estimate for the bilinear form $\mathcal{A}_0(\cdot,\cdot)$ in the energy norm $|||\cdot|||_{DG0}^2$ [see (2.18)]. For all $z \in \mathcal{Z}_\beta$ we have

$$\begin{aligned}\mathcal{A}_0(z,z) \ &\gtrsim \ |||z|||_{DG0}^2 = \sum_{T\in\mathcal{T}_h} \kappa_T\|\nabla z\|_{0,T}^2 + \sum_{e\in\mathcal{E}_h} \kappa_e\|h_e^{-1/2}\mathcal{P}_e^0([\![z]\!])\|_{0,e}^2 \\ &\gtrsim \ \sum_{e\in\mathcal{E}_h} \kappa_e\|h_e^{-1/2}\mathcal{P}_e^0([\![z]\!])\|_{0,e}^2 = \|z\|_*^2.\end{aligned}$$

Since this is the desired bound we conclude the proof of (4.5).

The proof of (4.6) follows easily from the spectral equivalence between $\mathcal{A}(\cdot,\cdot)$ and $\mathcal{A}_0(\cdot,\cdot)$ given in Lemma 2.2. ∎

Observe that last result guarantees that the linear systems on $\mathcal{Z}_\beta$ can be efficiently solved by preconditioned CG (PCG) with a diagonal preconditioner. As a corollary of the result in Lemma 4.3, the number of PCG iterations will be independent of both the mesh size and the variations in the PDE coefficient.

We now show that in the particular case of IIPG-0 method, the matrix representation of $\mathcal{A}_0(\cdot,\cdot)$ on $\mathcal{Z}_\beta$ is itself a diagonal matrix.

Lemma 4.4. *Let $\mathcal{A}_0(\cdot,\cdot)$ be the bilinear form of the non-symmetric IIPG(β) Type-0 method (2.9) with $\theta = 0$. Let $\{\psi_e^z\}_{e\in\mathcal{E}_h}$ be the basis for the space $\mathcal{Z}_\beta$ as defined in (3.5). Let $\mathbb{A}_0^{zz}$ be the matrix representation in this basis of the restriction to the subspace $\mathcal{Z}_\beta$ of the operator associated to $\mathcal{A}_0(\cdot,\cdot)$. Then, $\mathbb{A}_0^{zz}$ is diagonal.*

Proof. Note that from the definition (2.9) of the method ($\theta = 0$) together with (4.3) we have

$$\begin{aligned}\mathcal{A}_0(z,\psi) &= (\nabla z, \nabla\psi)_{\mathcal{T}_h} - \langle \{\!\{\nabla z\}\!\}_{\beta_e}, [\![\psi]\!]\rangle_{\mathcal{E}_h} + \langle \alpha_e h_e^{-1}\kappa_e \mathcal{P}_e^0([\![z]\!]), \mathcal{P}_e^0([\![\psi]\!])\rangle_{\mathcal{E}_h} \\ &= \langle \alpha_e h_e^{-1}\kappa_e \mathcal{P}_e^0([\![z]\!]), \mathcal{P}_e^0([\![\psi]\!])\rangle_{\mathcal{E}_h} \quad \forall\, z,\psi \in \mathcal{Z}_\beta\ . \end{aligned} \tag{4.9}$$

Let $\{\psi_e^z\}_{\{e\in\mathcal{E}_h\}}$ be the basis functions (3.5). To prove that $\mathbb{A}_0^{zz}$ is diagonal it is enough to show that for the basis functions (3.5), the following relation holds:

$$\mathcal{A}_0(\psi_e^z, \psi_{e'}^z) = c_e\delta_{e,e'} \qquad c_e \neq 0, \qquad \forall\, e \in \mathcal{E}_h\ , \tag{4.10}$$

where $\delta_{e,e'}$ is the delta function associated with the edge/face e. We now show (4.10). Observe that the supports of ψ_e^z and $\psi_{e'}^z$ have empty intersection unless $e, e' \subset T$ for some $T \in \mathcal{T}_h$. Let $T \cap \partial\Omega = \emptyset$ be an interior element, then from (4.9) and the mid-point integration rule, we have

$$\begin{aligned}\mathcal{A}_0(\psi_e^z, \psi_{e'}^z) &= \alpha_e h_e^{-1}\int_e \kappa_e \mathcal{P}_e^0([\![\psi_e^z]\!])\mathcal{P}_e^0([\![\psi_{e'}^z]\!]) = \alpha_e h_e^{-1}\kappa_e [2\psi_e^z(m_e)][2\psi_{e'}^z(m_e)] \\ &= 4\alpha_e h_e^{-1}\kappa_e\delta_{e,e'} \qquad e,e' \subset \partial T \quad e,e' \in \mathcal{E}_h^o\ , \end{aligned}$$

which shows (4.10) for interior edges with $c_e = 4\alpha_e h_e^{-1}\kappa_e$. For boundary edges/faces the considerations are essentially the same and therefore omitted. The proof is complete since the relation (4.10) readily implies that the off-diagonal terms of $\mathbb{A}_0^{zz}$ are zero. □

We remark that this lemma will be essential for proving the uniform convergence of the iterative solvers for the non-symmetric methods.

5. Robust preconditioners on V_h^{CR}

In this section we develop efficient and robust solvers for the solution in the CR space. By virtue of the spectral equivalence between $\mathcal{A}$ and $\mathcal{A}_0$ given in Lemma 2.2, it is enough to focus on the development of solvers for $\mathcal{A}_0$ restricted to the subspace V_h^{CR} and this is what we do next. Observe that the restriction of $\mathcal{A}_0(\cdot,\cdot)$ to the subspace V^{CR} reduces to the $\mathbb{P}^1$-nonconforming finite element discretization of (1.1)

$$\text{Find } u \in V_h^{CR}:\ \ \mathcal{A}_0(u,w) = (\kappa\nabla u, \nabla w)_{\mathcal{T}_h} = (f,w) \quad \forall\, w \in V_h^{CR}. \tag{5.1}$$

We denote A_0^{CR} as the operator induced by (5.1). For the analysis that follows, we need the following semi-norms and norms for any $v \in V_h^{CR}$:

$$|v|_{1,h,\kappa}^2 := \sum_{T\in\mathcal{T}_h} \kappa_T \|\nabla v\|_{0,T}^2\ , \qquad |v|_{1,h,\Omega_i}^2 := \sum_{T\in\mathcal{T}_h\ ,T\subseteq\Omega_i} \|\nabla v\|_{0,T}^2\ , \tag{5.2}$$

$$\|v\|_{0,\kappa}^2 := \sum_{i=1}^M \kappa\big|_{\Omega_i}\|v\|_{0,\Omega_i}^2\ , \qquad \|v\|_{1,h,\kappa}^2 := \|v\|_{0,\kappa}^2 + |v|_{1,h,\kappa}^2\ . \tag{5.3}$$

If the coefficient $\kappa(x)$ is a constant or a smooth function; one can find some uniform preconditioners in the literature (see the references in [9]). The case of jump coefficients κ has also

been considered by several authors [65], [64], with the assumption on the quasi-uniformity in the coefficients.

Since (5.1) is a symmetric problem, from the classical theory of PCG we know that the convergence rates of the iterative method with preconditioner, say B, for A_0^{CR} are fully determined, in the worst case scenario, by the condition number of the preconditioned system: $\mathcal{K}(BA_0^{CR})$. However, for problems with large jumps in the coefficient κ, it has also been observed in [49, 78] that the spectrum $\sigma(BA_0^{CR})$ might contain a few very small eigenvalues, which result in a extremely large value of $\mathcal{K}(BA_0^{CR})$ (independently of how fine is the mesh), but they seem to have no influence on the efficiency of the preconditioner and the overall convergence of the iterative method. Therefore, following the approach introduced in [78], we consider the *asymptotic convergence rate* of the PCG algorithm, which is determined by the so-called *effective condition number*, as defined below in Definition 5.1.

The setting that we need for the analysis require some more preliminaries. Let us first briefly review few standard results for the PCG algorithm. Suppose that the spectrum of BA_0^{CR}, $\sigma(BA_0^{CR})$, is divided in two sets: $\sigma(BA_0^{CR}) = \sigma_0(BA_0^{CR}) \cup \sigma_1(BA_0^{CR})$, where $\sigma_0(BA_0^{CR}) = \{\lambda_1, \ldots, \lambda_{m_0}\}$ contains of all very small (often referred to as "bad") eigenvalues and the remaining eigenvalues (bounded above and below) are in $\sigma_1(BA_0^{CR}) = \{\lambda_{m_0+1}, \ldots, \lambda_N\}$; that is $\lambda_j \in [a, b]$ for $j = m_0 + 1, \ldots, N$, with $N = \dim(V_h^{CR}) = n_E - n_{BE}$. Then, the error at the k-th iterate of the PCG algorithm is bounded by (see e.g. [6, 51, 7]):

$$\|u - u_k\|_{1,h,\kappa} \leq 2(\mathcal{K}(BA_0^{CR}) - 1)^{m_0} \left(\frac{\sqrt{b/a} - 1}{\sqrt{b/a} + 1} \right)^{k-m_0} \|u - u_0\|_{1,h,\kappa} \, . \tag{5.4}$$

From the above estimate, one may conclude the following: If there are only a few small eigenvalues of BA_0^{CR} in $\sigma_0(BA_0^{CR})$, then the *asymptotic* convergence rate of the resulting PCG method will be dominated by the factor $\frac{\sqrt{b/a}-1}{\sqrt{b/a}+1}$, i.e. by b/a where $b = \lambda_N(BA_0^{CR})$ and $a = \lambda_{m_0+1}(BA_0^{CR})$. The quantity (b/a) which determines the asymptotic convergence rate is often called *effective condition number* and its formal definition follows.

Definition 5.1. Let V be a real N-dimensional Hilbert space, and $\mathfrak{A} : V \to V$ be a symmetric positive definite linear operator, with eigenvalues $0 < \lambda_1 \leq \cdots \leq \lambda_N$. The *$m_0$-th effective condition number* of $\mathfrak{A}$ is defined by

$$\mathcal{K}_{m_0+1}(\mathfrak{A}) := \frac{\lambda_N(\mathfrak{A})}{\lambda_{m_0+1}(\mathfrak{A})}.$$

We note that this characteristic was previously used in the context of deflation based preconditioner (for determining the asymptotic convergence rate of PCG for semidefinite systems [45]).

In our analysis of two-level and multilevel preconditioners for problem (5.1), we will mainly focus on estimating the effective condition number. Related works which derive estimates on the effective condition number in the context of multigrid and domain decomposition methods for *conforming* discretizations are [49, 78, 80]. To improve the flow of the discussion when presenting the analysis of the multilevel preconditioner for the nonconforming discretizations, we start by introducing a two level method and provide a detailed analysis and bounds of the condition number of the preconditioned system.

5.1. **Two level preconditioner for $\mathcal{A}_0(\cdot,\cdot)$ on V_h^{CR}.** In this subsection we will construct an additive preconditioner (parallel subspace correction method) with a conforming space $V_{\widetilde{h}}^{\text{conf}} \subset V_h^{CR}$ as a coarse space plus a pointwise relaxation (point-Jacobi or Gauss-Seidel method). More precisely to define the two level parallel subspace correction method, we consider the following

overlapping space decomposition of V_h^{CR}:

$$V_h^{CR} = V_h^{CR} + V_{\widetilde{h}}^{\text{conf}}. \tag{5.5}$$

Here, we allow for different $\widetilde{h}$: one could set $\widetilde{h} = h$, or choose a coarser approximation i.e., $\widetilde{h} = H > h$. In addition we will assume here that the meshes $\mathcal{T}_h$ and $\mathcal{T}_{\widetilde{h}}$ are nested. On $V_{\widetilde{h}}^{\text{conf}}$ we consider the standard conforming $\mathbb{P}^1$-approximation to (1.1): Find $\chi \in V_{\widetilde{h}}^{\text{conf}}$ such that

$$a(\chi, \eta) := \mathcal{A}_0(\chi, \eta) = \int_\Omega \kappa \nabla\chi \cdot \nabla\eta dx = (f, \eta), \quad \forall\, \eta \in V_{\widetilde{h}}^{\text{conf}}\,. \tag{5.6}$$

The bilinear form given above defines an "energy" scalar product and its naturally induced weighted semi-norm is:

$$|\chi|^2_{1,\kappa,D} := \int_D \kappa |\nabla\chi|^2 dx\,, \quad \forall\, \chi \in H^1(D), \quad D \subset \Omega. \tag{5.7}$$

We next discuss the matrix form of the additive preconditioner and then its operator form. The reason for discussing both these *equivalent* forms is that the matrix form is used to implement the method, while the operator form is more suitable for the analysis and the spectral equivalence results that we prove.

5.1.1. ***Definition of the preconditioner: Matrix notation.*** We begin by introducing the necessary notation pertaining to vectors and matrices. We write $\mathbb{A}_0^{CR} = \mathbb{A}_0^{vv} = \mathbb{D}_0 - \mathbb{L} - \mathbb{L}^t$ where $\mathbb{D}_0$, and $\mathbb{L}$ are the diagonal and the strict lower triangular part of the stiffness matrix $\mathbb{A}_0^{CR}$ representing the restriction of $\mathcal{A}_0(\cdot,\cdot)$ on V_h^{CR}. We use similar notation for the space $V_{\widetilde{h}}^{\text{conf}}$ denoting by $\mathbb{A}^C$ the stiffness matrix representing the bilinear form $a(\cdot,\cdot)$ on the conforming space $V_{\widetilde{h}}^{\text{conf}}$. For $v \in V_h^{CR}$, $v(x) = \sum_{e\in\mathcal{E}_h^o} v_e \varphi_e^{CR}(x)$, we denote by $\mathbf{v} \in \mathbb{R}^{n_{CR}}$ the vector of degrees of freedom of $v(x)$, that is $\mathbf{v}$ has components $\{v_e\}_{e\in\mathcal{E}_h^o}$. Here $n_{CR} = n_E - n_{BE}$ is the dimension of the space V_h^{CR} and we shall denote by n_C the dimension of $V_{\widetilde{h}}^{\text{conf}}$. The latter dimension is equal to the number of interior vertices in the triangulation with mesh size $\widetilde{h}$. We then have the following relation between stiffness matrices and bilinear forms and vector representations of degrees of freedom:

$$(\mathbb{A}_0^{CR}\mathbf{v}, \mathbf{w})_{\ell_2} = \mathcal{A}_0(v, w), \qquad (\mathbb{A}^C\boldsymbol{\chi}, \boldsymbol{\eta})_{\ell_2} = a(\chi, \eta). \tag{5.8}$$

These identities hold for all $v \in V_h^{CR}$ and $w \in V_h^{CR}$ and their vectors of degrees of freedom, $\mathbf{v} \in \mathbb{R}^{n_{CR}}$ and $\mathbf{w} \in \mathbb{R}^{n_{CR}}$, as well as for all $\chi \in V_{\widetilde{h}}^{\text{conf}}$, and all $\eta \in V_{\widetilde{h}}^{\text{conf}}$ and the respective vectors of degrees of freedom $\boldsymbol{\chi} \in \mathbb{R}^{n_C}$ and $\boldsymbol{\eta} \in \mathbb{R}^{n_C}$.

Since the conforming finite element functions are in V_h^{CR}, we may expand each of the canonical basis functions from $V_{\widetilde{h}}^{\text{conf}}$ via the basis in V_h^{CR}. The coefficients in such expansions form a matrix $\Pi \in \mathbb{R}^{n_{CR}\times n_c}$ which represents the *natural inclusion* operator $V_{\widetilde{h}}^{\text{conf}} \subset V_h^{CR}$. The additive preconditioner $\mathbb{B} : \mathbb{R}^{n_{CR}} \mapsto \mathbb{R}^{n_{CR}}$ is then defined in the usual manner:

$$\mathbb{B} := \mathbb{D}_0^{-1} + \Pi(\mathbb{A}^C)^{-1}\Pi^t. \tag{5.9}$$

Denoting by $\mathcal{N}_0(\mathcal{T}_{\widetilde{h}})$ the set of interior vertices of $\mathcal{T}_{\widetilde{h}}$ we see that the following relation holds for all $p \in \mathcal{N}_0(\mathcal{T}_{\widetilde{h}})$ and $p' \in \mathcal{N}_0(\mathcal{T}_{\widetilde{h}})$:

$$[\mathbb{A}^C]_{p,p'} = a(\varphi_p, \varphi_{p'}) = \left[\Pi^T\mathbb{A}^{CR}\Pi\right]_{p,p'}, \quad \text{i.e.} \quad \mathbb{A}^C = \Pi^T\mathbb{A}^{CR}\Pi. \tag{5.10}$$

5.1.2. ***Definition of the preconditioner: Operator notation.*** We first consider the operator R^{-1} corresponding to the Jacobi smoother. We note that the usual definition of the operator associated with Jacobi relaxation requires additional notation, including introducing mass matrices and other objects, which are not actually used in the implementation of the algorithms. To avoid introducing these extra quantities, we will take a slightly different path and we define $R^{-1}A_0^{CR}$ instead of R^{-1}. We set:

$$R^{-1}A_0^{CR}v = \sum_{e\in\mathcal{E}_h^o} \frac{\mathcal{A}_0(v,\varphi_e^{CR})}{\mathcal{A}_0(\varphi_e^{CR},\varphi_e^{CR})}\,\varphi_e^{CR} = \sum_e \frac{[\mathbb{A}_0^{CR}\mathbf{v}]_e}{[\mathbb{A}_0^{CR}]_{e,e}}\,\varphi_e^{CR} \tag{5.11}$$

It is clear from this definition that the vector of degrees of freedom for the function $R^{-1}A_0^{CR}v$ is $\mathbb{D}_0^{-1}\mathbb{A}_0^{CR}\mathbf{v}$, which precisely shows that the matrix representation of R^{-1} is $\mathbb{D}_0^{-1}$.

We next consider the operator corresponding to the correction from the conforming space $V_{\widetilde{h}}^{\text{conf}}$. For its definition we need to introduce the standard $\mathcal{A}_0(\cdot,\cdot)$-orthogonal and $L_2(\Omega)$-orthogonal projections on $V_{\widetilde{h}}^{\text{conf}}$. Let us denote these projections with P^C and Q^C, respectively. They are defined for any $v\in V_h^{CR}$ and $[P^Cv]\in V_{\widetilde{h}}^{\text{conf}}$ and $[Q^Cv]\in V_{\widetilde{h}}^{\text{conf}}$ are as follows:

$$a([P^Cv],\eta) = \mathcal{A}_0(v,\eta), \quad ([Q^Cv],\eta) = (v,\eta), \qquad \forall \quad \eta\in V_{\widetilde{h}}^{\text{conf}}. \tag{5.12}$$

The operator form of the preconditioner is then

$$B: V_h^{CR}\mapsto V_h^{CR}, \quad B := R^{-1} + (A^C)^{-1}Q^C. \tag{5.13}$$

We also have the following identity, which follows directly from the definition of Q^C and P^C:

$$Q^CA_0^{CR} = A^CP^C. \tag{5.14}$$

Multiplying from the right the defining relation for B in (5.13) with A_0^{CR} and using the above relation (5.14) we have that

$$BA_0^{CR} := R^{-1}A_0^{CR} + P^C. \tag{5.15}$$

From the definitions given above it is immediate to show that the vector of degrees of freedom corresponding to the function $BA_0^{CR}v$ is $\mathbb{B}\mathbb{A}_0^{CR}\mathbf{v}$, where $\mathbf{v}$ is the vector of degrees of freedom for $v(x)$. This together with (5.8) shows that the following identity holds for any $v\in V_h^{CR}$:

$$\frac{(\mathbb{B}\mathbb{A}_0^{CR}\mathbf{v},\mathbb{A}_0^{CR}\mathbf{v})_{\ell_2}}{(\mathbb{A}_0^{CR}\mathbf{v},\mathbf{v})_{\ell_2}} = \frac{\mathcal{A}_0(BA_0^{CR}v,v)}{\mathcal{A}_0(v,v)} \tag{5.16}$$

Note that the left side of this identity uses the ℓ_2 inner product on $\mathbb{R}^{n_{CR}}$, while calculating the right side is done through the bilinear form $\mathcal{A}_0(\cdot,\cdot)$.

5.1.3. ***Spectral equivalence results for the two level preconditioner.*** We recall definition of the set of indices $\mathcal{I}$ of *floating* subdomains (see [69]) which are the subdomains where the coefficient is constant and which do not touch the Dirichlet boundary:

$$\mathcal{I} := \{\ i\ :\ \text{meas}_{d-1}(\partial\Omega\cap\partial\Omega_i) = 0\ \}\ . \tag{5.17}$$

The role of the set $\mathcal{I}$ will be clear from the convergence analysis given later on in this section.

The main spectral equivalence result for the two level preconditioner is as follows.

Theorem 5.2. *Let $m_0 = |\mathcal{I}|$ be the total number of floating subdomains. Then, the effective condition number $\mathcal{K}_{eff}(BA_0^{CR}) := \mathcal{K}_{m_0+1}(BA_0^{CR})$ for the preconditioner B defined in* (5.13) *is uniformly bounded:*

$$\mathcal{K}_{m_0+1}(BA_0) \lesssim (1+|\log 2\widetilde{h}/h|) \le C, \qquad \text{for } \widetilde{h}\ge h\ ,$$

with $C>0$ independent of the coefficients.

The rest of the section is devoted to the proof of Theorem 5.2, which essentially involves showing spectral equivalence of the form

$$\mathcal{A}_0(v,v) \lesssim (B^{-1}v,v) \lesssim \mathcal{A}_0(v,v).$$

It can be shown (see Appendix B, Lemma B.2) that the inverse of B satisfies:

$$(B^{-1}v,v) = \inf_{\chi \in V_{\tilde{h}}^{\mathrm{conf}}} [\mathcal{R}(v-\chi, v-\chi) + a(\chi,\chi)], \tag{5.18}$$

where $\mathcal{R}(\cdot,\cdot)$ is the bilinear form associated with the smoother and is defined as: $\mathcal{R}(v,w) := (Rv,w)$. In the Appendix B (see Lemma B.1) it will be shown that for all $v \in V_h^{CR}$ we have

$$\mathcal{R}(v,v) = \sum_{e \in \mathcal{E}_h^o} \mathcal{A}_0(\varphi_e^{CR}, \varphi_e^{CR}) v_e^2 (\mathbb{D}_0 \mathbf{v}, \mathbf{v})_{\ell_2}. \tag{5.19}$$

From the above considerations it is clear that the proof of Theorem 5.2 amounts to verifying the following two conditions.

(F1) There exists a constant c_1 such that for all $v \in V_h^{CR}$ we have

$$c_1^{-1}\mathcal{A}_0(v,v) \lesssim \mathcal{R}(v,v).$$

(F2) For every $v \in V_h^{CR}$ there exists $\chi_v \in V_{\tilde{h}}^{\mathrm{conf}}$ (depending on v), such that

$$\mathcal{R}(v-\chi_v, v-\chi_v) + a(\chi_v,\chi_v) \lesssim c_0 \mathcal{A}_0(v,v).$$

We remark here that the identity (5.18) is a well known result regarding two level additive Schwarz methods (cf. [77, 67, 69]).

The validation of last assumption **(F2)** is more intricate and we postpone it to the next subsection, while the Lemma stated next shows that **(F1)** is true for our choice of $\mathcal{R}(\cdot,\cdot)$, and also gives an inequality which will later be used to verify **(F2)**.

Lemma 5.3. *Let $\mathcal{R}(\cdot,\cdot)$ be the bilinear form defined via* (5.19)*. Then we have the following estimates*

$$c_1^{-1}\mathcal{A}_0(v,v) \lesssim \mathcal{R}(v,v) \quad \textit{and} \quad \mathcal{R}(v,v) \simeq h^{-2}\|v\|_{0,\kappa}^2\,, \quad \forall\, v \in V_h^{CR}\,. \tag{5.20}$$

Proof. To show that **(F1)** holds we set $w = \sum_e w_e \varphi_e^{CR}$ and Cauchy-Schwarz and the arithmetic-geometric inequalities give

$$\begin{aligned}
\mathcal{A}_0(v,v) &= \sum_e \sum_{e'} \mathcal{A}_0(\varphi_e^{CR}, \varphi_{e'}^{CR}) v_e v_{e'} \\
&\le \sum_e \sum_{e'} \sqrt{\mathcal{A}_0(\varphi_e^{CR}, \varphi_e^{CR})}\sqrt{\mathcal{A}_0(\varphi_{e'}^{CR}, \varphi_{e'}^{CR})} v_e v_{e'} \\
&\le \frac{1}{2}\sum_e \sum_{e'} \left[\mathcal{A}_0(\varphi_e^{CR}, \varphi_e^{CR}) v_e^2 + \mathcal{A}_0(\varphi_{e'}^{CR}, \varphi_{e'}^{CR}) v_{e'}^2\right] \\
&= \sum_e \sum_{e'} \mathcal{A}_0(\varphi_e^{CR}, \varphi_e^{CR}) v_e^2 \lesssim \sum_e \mathcal{A}_0(\varphi_e^{CR}, \varphi_e^{CR}) v_e^2 = \mathcal{R}(v,v).
\end{aligned}$$

The constant hidden in the "$\lesssim$" above only depends on the number of neighboring faces e', for a given face $e \in \mathcal{E}_h^o$ and this constant is bounded by 5 in 2D and 7 in 3D.

Notice that since the mesh is quasi-uniform, for any $V_h^{CR} \ni v = \sum_e v_e \varphi_e^{CR}$ and $T \in \mathcal{T}_h$ we have

$$\|v\|_{0,\kappa,T}^2 \simeq \sum_{e \subset \partial T} v_e^2 \|\varphi_e^{CR}\|_{0,\kappa,T}^2. \tag{5.21}$$

On one hand, for any basis function φ_e^{CR} by the inverse inequality (see [29]) we have that: $h^{-2}\|\varphi_e^{CR}\|^2_{0,\kappa,T} \lesssim \|\varphi_e^{CR}\|^2_{0,\kappa,T}$. On the other hand, directly calculating the norms gives us the following bound $\|\nabla\varphi_e^{CR}\|^2_{0,\kappa,T} \lesssim h^{-2}\|\varphi_e^{CR}\|^2_{0,\kappa,T}$. Hence,

$$h^{-2}\|\varphi_e^{CR}\|^2_{0,\kappa,T} \simeq \|\nabla\varphi_e^{CR}\|^2_{0,\kappa,T}. \tag{5.22}$$

Therefore from the definition of $\mathcal{R}(\cdot,\cdot)$ given in (5.19) and the equivalence relations (5.21) and (5.22) we get

$$\begin{aligned}
\mathcal{R}(v,v) &= \sum_e v_e^2\|\nabla\varphi_e^{CR}\|^2_{0,\kappa} = \sum_e v_e^2 \sum_{\partial T\supset e} \|\nabla\varphi_e^{CR}\|^2_{0,\kappa,T} \\
&= \sum_{T\in\mathcal{T}_h}\sum_{e\subset\partial T} v_e^2\|\nabla\varphi_e^{CR}\|^2_{0,\kappa,T} \simeq \sum_{T\in\mathcal{T}_h}\sum_{e\subset\partial T} h^{-2}v_e^2\|\varphi_e^{CR}\|^2_{0,\kappa,T} \\
&\simeq h^{-2}\sum_{T\in\mathcal{T}_h}\|v\|^2_{0,\kappa,T} = h^{-2}\|v\|^2_{0,\kappa}.
\end{aligned}$$

□

5.2. **A stable Decomposition.** In this subsection we show that condition **(F2)** is satisfied and the main tool is an operator $P_h^{\widetilde{h}} : V_h^{CR} \to V_{\widetilde{h}}^{\text{conf}}$, that will be shown to satisfy certain approximation and stability properties as stated in the next Lemma.

Lemma 5.4. *There exists an interpolation operator $P_h^{\widetilde{h}} : V_h^{CR} \to V_{\widetilde{h}}^{\text{conf}}$, that satisfies the following approximation and stability properties:*

$$\textit{Approximation:} \qquad \|(I-P_h^{\widetilde{h}})v\|_{0,\kappa} \le C_a\widetilde{h}|\log 2\widetilde{h}/h|^{1/2}\|v\|_{1,h,\kappa}, \qquad \forall\, v\in V_h^{CR}, \tag{5.23}$$

$$\textit{Stability:} \qquad |P_h^{\widetilde{h}}v|_{1,\kappa} \le C_s|\log 2\widetilde{h}/h|^{1/2}\|v\|_{1,h,\kappa} \qquad \forall\, v\in V_h^{CR}, \tag{5.24}$$

with constants C_a and C_s independent of the coefficient κ and mesh size.

The definition of such $P_h^{\widetilde{h}}$ and the proof of the above result is given in the Appendix A. We would like to point out that the operator $P_h^{\widetilde{h}}$ is not used in the actual implementation of the preconditioner $\mathbb{B}$, as is plainly seen from (5.9). However, the operator $P_h^{\widetilde{h}}$ and its approximation and stability properties play crucial role in the analysis.

We now discuss and display the main ingredients in the proof of condition **(F2)**, with the aid of Lemma (5.4). Observe that on the right hand side of (5.23) and (5.24), the bounds are given in terms of the weighted full H^1-norm $\|v\|_{1,h,\kappa}$. In general, one cannot replace the norm $\|v\|_{1,h,\kappa}$ by the energy semi-norm $|v|_{1,h,\kappa}$, unless some stronger conditions on the coefficients (like quasi-monotonicity) are imposed. To be able to replace the full norm by the semi-norm, we introduce the subspace $\widetilde{V}_h^{CR} \subset V_h^{CR}$:

$$\widetilde{V}_h^{CR} := \left\{ v\in V_h^{CR} : \int_{\Omega_i} v dx = 0, \;\; \forall i\in\mathcal{I} \right\}, \tag{5.25}$$

where $\mathcal{I}$ is the set of indexes of the *floating* subdomains as defined in (5.17). Recall that a *floating* subdomain is a subdomain whose boundary does not intersect the Dirichlet boundary. The key feature of the subspace we have just introduced is that the Poincaré-Friedrichs inequality for the nonconforming finite element space [36, 17] holds on each subdomain, and this allows us to replace the full norm $\|v\|_{1,h,\kappa}$ by the energy semi-norm $|v|_{1,h,\kappa}$, for any $v\in\widetilde{V}_h^{CR}$. We should remark that the condition on the average, i.e. $\int_{\Omega_i} v dx = 0$, is not essential; other type of conditions could be used (see [69]) as long as they allow for the application of a Poincarè-type inequality.

At this point we would like to mention that the dimension of $\widetilde{V}_h^{CR}$ is related to the number of floating subdomains and in fact, $\dim(\widetilde{V}_h^{CR}) = \dim(V_h^{CR}) - m_0$. By restricting now the action of the operator $P_h^{\widetilde{h}}$ to functions in $\widetilde{V}_h^{CR}$, we have the following result, as an easy corollary from Lemma 5.4.

Corollary 5.5. *Let $\widetilde{V}_h^{CR} \subset V_h^{CR}$ the subspace of the Crouziex-Raviart finite element space, as defined in* (5.25)*. Then, there exist an operator $P_h^{\widetilde{h}} : V_h^{CR} \to V_{\widetilde{h}}^{\text{conf}}$ that is stable in the weighted H^1-broken semi-norm:*

$$|P_h^{\widetilde{h}} v|_{1,\kappa} \lesssim |\log 2\widetilde{h}/h|^{1/2} |v|_{1,h,\kappa} , \quad \forall\, v \in \widetilde{V}_h^{CR} ,$$

and satisfies the following approximation property:

$$\|(I - P_h^{\widetilde{h}})v\|_{0,\kappa} \lesssim \widetilde{h} |\log 2\widetilde{h}/h|^{1/2} |v|_{1,h,\kappa} , \quad \forall\, v \in \widetilde{V}_h^{CR} .$$

Proof. The proof follows from Lemma 5.4 together with a standard application of the Poincarè-Friedrichs type inequalities inequality for nonconforming functions, which allows for bounding the norm $\|v\|_{1,h,\kappa}$ by the semi-norm $|v|_{1,h,\kappa}$. In fact, from the definition of $\|v\|_{1,h,\kappa}$:

$$\begin{aligned} \|v\|_{1,h,\kappa}^2 = \sum_{\Omega_i \subseteq \Omega} \|v\|_{1,h,\kappa,\Omega_i}^2 \quad &= \quad \sum_{i \notin \mathcal{I}} \left(\kappa \|v\|_{0,\Omega_i}^2 + \kappa |v|_{1,h,\Omega_i}^2 \right) \\ &\quad + \sum_{i \in \mathcal{I}} \left(\kappa \|v\|_{0,\Omega_i}^2 + \kappa |v|_{1,h,\Omega_i}^2 \right) . \end{aligned}$$

For the terms in both sums above, one can apply Poincarè-Friedrichs type inequalities. In the first sum above, for $i \notin \mathcal{I}$ each subdomain touches the boundary, and hence $\|v\|_{0,\Omega_i} \leq C_p |v|_{1,h,\Omega_i}$. For the terms in the second sum, since $v \in \widetilde{V}_h^{CR}$ are functions with average zero on Ω_i, we have $\|v\|_{0,\Omega_i} \leq C_p |v|_{1,h,\Omega_i}$, $i \in \mathcal{I}$. □

Remark 5.6. In particular, if we consider the case when the ratio $\widetilde{h}/h$ is a fixed constant (e.g. $\widetilde{h} = h$) then the approximation and stability properties are:

$$\begin{aligned} \|(I - P_h^{\widetilde{h}})v\|_{0,\kappa} \leq C \log(2) h |v|_{1,h,\kappa} \lesssim h |v|_{1,h,\kappa}, \\ |P_h^{\widetilde{h}} v|_{1,\kappa} \leq C \log(2) |v|_{1,h,\kappa} \lesssim |v|_{1,h,\kappa}. \end{aligned}$$

With the aid of the result in Corollary 5.5, we are able to show that the stability condition **(F2)** holds for functions from the subspace $\widetilde{V}_h^{CR}$.

Lemma 5.7. *For any $v \in \widetilde{V}_h^{CR}$ the stability condition **(F2)** holds, namely, there exist a constant $c_0 > 0$ and a function $\chi = P_h^{\widetilde{h}}(v) \in V_{\widetilde{h}}^{\text{conf}}$ such that*

$$\mathcal{R}(v - \chi, v - \chi) + a(\chi, \chi) \leq c_0 |v|_{1,h,\kappa}^2, \qquad c_0 = C(1 + \log \frac{2\widetilde{h}}{h}) . \tag{5.26}$$

Proof. Given any $v \in \widetilde{V}_h^{CR}$, let $\chi \in V_{\widetilde{h}}^{\text{conf}}$ be defined as $\chi := P_h^{\widetilde{h}} v$. Taking into account the approximation property of $P_h^{\widetilde{h}}$ (with $\widetilde{h} = h$) given in (5.5) together with Remark 5.6, we have

$$\mathcal{R}(v - \chi, v - \chi) \lesssim h^{-2} \|v - \chi\|_{0,\kappa}^2 \lesssim h^{-2} \|v - P_h^{\widetilde{h}} v\|_{0,\kappa}^2 \lesssim |v|_{1,h,\kappa}^2 = \mathcal{A}_0(v, v).$$

The proof is now complete by using the stability of $P_h^{\widetilde{h}} v$ given in Corollary 5.5;

$$a(\chi, \chi) = |P_h^{\widetilde{h}} v|_{1,\kappa}^2 \lesssim |\log 2\widetilde{h}/h| |v|_{1,h,\kappa}^2 = |\log 2\widetilde{h}/h| \mathcal{A}_0(v, v).$$

□

We have now all ingredients to complete the proof of Theorem 5.2.

Proof of Theorem 5.2. To estimate the maximum eigenvalue of BA_0^{CR} let $\chi \in V_h^C$ and $v \in V_h^{CR}$ be arbitrary. We set $v_0 = (v - \chi)$, and hence $v = v_0 + \chi$. Applying the Cauchy-Schwarz inequality, the arithmetic-geometric inequality, and (5.20) then yields

$$\begin{aligned}\mathcal{A}_0(v,v) &= \mathcal{A}_0(v_0+\chi, v_0+\chi) = \mathcal{A}_0(v_0,v_0) + \mathcal{A}_0(\chi,\chi) + 2\mathcal{A}_0(v_0,\chi)\\ &\le 2\mathcal{A}_0(v_0,v_0) + 2\mathcal{A}_0(\chi,\chi) \le c_1\left(\mathcal{R}(v_0,v_0) + a(\chi,\chi)\right),\end{aligned}$$

Since $\chi \in V_{\widetilde{h}}^{\text{conf}}$ was arbitrary, taking the infimum with respect to χ and using the definition of $(B^{-1}v, v)$ (5.18) shows that

$$\mathcal{A}_0(v,v) \lesssim (B^{-1}v, v).$$

This guarantees that the maximum eigenvalue $\lambda_{\max}(BA_0^{CR}) \le c_1$ is uniformly bounded independently of the coefficient and the mesh size.

For the lower bound, by Lemma 5.7 we have that for any $v \in \widetilde{V}_h^{CR}$ with $\chi_v = P_h^{\widetilde{h}} v$ we have

$$\mathcal{R}(v - \chi_v, v - \chi_v) + |\chi_v|_{1,\kappa}^2 \le c_0 |v|_{1,h,\kappa}^2 = c_0 \mathcal{A}_0(v,v).$$

Then by the *minimax principle* (cf. [52, §90], [6, Lemma 3.13] or [46, Theorem 8.1.2]) and noticing that $\dim(\widetilde{V}_h^{CR}) = \dim(V_h^{CR}) - m_0$, we have $\lambda_{m_0+1} \ge c_0^{-1}$. Therefore, the effective condition $\mathcal{K}_{m_0+1}(BA_0^{CR})$ can be bounded by

$$\mathcal{K}_{m_0+1}(BA_0^{CR}) \le c_0 c_1,$$

with c_0 given in (5.26). This completes the proof. □

Remark 5.8. All the results that we have proved above also hold for R^{-1} defined by the symmetric Gauss-Seidel method. This follows from the fact that the bilinear forms $\mathcal{R}(\cdot,\cdot)$ defined via Jacobi or symmetric Gauss-Seidel method are equivalent with constants independent of the variations in the PDE coefficient. For a proof of such equivalence we refer to [71, Proposition 6.12] or [81, Lemma 3.3]. Thus, the two level (and also multilevel) preconditioners that use Jacobi or Gauss-Seidel as smoother are equivalent and hence all the results here stated for the Jacobi smoother also hold for the symmetric Gauss-Seidel smoother.

5.3. **Multilevel Preconditioner for $\mathcal{A}_0(\cdot,\cdot)$ on V_h^{CR} .** We now introduce multilevel preconditioner, using the two level theory developed before. The multilevel preconditioner corresponds to replacing $[A^C]^{-1}$ in (5.13) with a spectrally equivalent operator $B^C : V_{\widetilde{h}}^{\text{conf}} \mapsto V_{\widetilde{h}}^{\text{conf}}$ corresponding to the additive BPX preconditioner (see e.g. [13, 14, 77]).

The space decomposition that we use to define the multilevel BPX preconditioner is:

$$V_h^{CR} = V_h^{CR} + \sum_{j=0}^{J} W_j = \sum_{j=0}^{J+1} W_j, \tag{5.27}$$

where we have denoted $W_j = V_{h_j}^{\text{conf}}$ $(j = 0, 1, \cdots, J)$ the nested conforming spaces with mesh sizes $h_j = 2^j h$ and $W_{J+1} = V_h^{CR}$. Let us also denote by A_j^C the operators corresponding to the restrictions of $a(\cdot,\cdot)$ on W_j for $j = 0, \cdots, J$, namely

$$(A_j^C v_j, w_j) = a(v_j, w_j), \quad \forall v_j \in W_j, \quad \forall w_j \in W_j.$$

Without loss of generality, we assume that $h \simeq 2^{-J}$, and therefore $|\log h| \simeq J$. Then the operator form of the multilevel preconditioner is:

$$B_{\text{ML}} : V_h^{CR} \mapsto V_h^{CR}, \quad B_{\text{ML}} := [A_0^C]^{-1} Q_0^C + \sum_{j=1}^{J+1} R_j^{-1} Q_j \tag{5.28}$$

Here $Q_j : V_h^{CR} \mapsto W_j$ is the L_2-orthogonal projection on W_j, $j = 0, \ldots, J$ and we set $Q_{J+1} = I$. We use exact solver on the coarsest grid. With all this notation in hand one can prove that

$$(B_{\mathrm{ML}}^{-1} v, v) = \inf_{\sum_{j=0}^{J+1} w_j = v} \left[a(w_0, w_0) + \sum_{j=1}^{J+1} \mathcal{R}_j(w_j, w_j) \right]. \tag{5.29}$$

Here $\mathcal{R}_j(\cdot,\cdot)$, $j = 1, \ldots, (J+1)$ correspond to Jacobi or symmetric Gauss-Seidel smoother and the proof of (5.29) is similar to the proof in the two level case.

Let us recall two results that we need later in our convergence analysis.

Lemma 5.9 ([78, Lemma 4.2])**.** *Let $\mathcal{R}_j(\cdot,\cdot)$ be a Jacobi or the symmetric Gauss-Siedel smoother for the solution of the discretization* (5.6) *on W_j space. Then,*

$$a(w,w) \lesssim \mathcal{R}_j(w,w) \lesssim h_j^{-2} \|w\|_{0,\kappa}^2 \quad \forall\, w \in W_j$$

We also need the following strengthened Cauchy Schwarz inequality.

Lemma 5.10 (Strengthened Cauchy Schwarz, cf. [77, Lemma 6.2])**.** *Let $j = 0, 1, \cdots, J-1$ and $j < l \leq J$ there exists a constant $\gamma \in (0,1)$ such that*

$$a(w_l, w_j) \lesssim \gamma^{l-j} (h_l^{-1} \|w_l\|_{0,\kappa})(h_j^{-1} \|w_j\|_{0,\kappa}), \quad \forall w_l \in W_l, \text{ and } w_j \in W_j. \tag{5.30}$$

The main result of this section is the following:

Theorem 5.11. *Let B_{ML} be the multilevel preconditioner defined in* (5.28) *and let m_0 denote the number of floating subdomains (the cardinality of the set $\mathcal{I}$ defined in* (5.17)*). Then, the following upper bound can be given for the effective condition number $\mathcal{K}_{m_0+1}(B_{ML} A_0^{CR})$ for the multilevel preconditioner B_{ML}:*

$$\mathcal{K}_{m_0+1}(B_{ML} A_0^{CR}) \leq C |\log h|^2 \lesssim |\log h|^2 \,,$$

where the constant $C > 0$ is independent of the coefficients and mesh size.

Proof. To estimate the effective condition number, we restrict our considerations to $v \in \widetilde{V}_h^{CR}$. We first give a bound on the *first relevant* eigenvalue $\lambda_{m_0+1}(B_{\mathrm{ML}}^{-1} A_0^{CR})$. From (5.29) we can see that to estimate this eigenvalue we need to find a decomposition $v = \sum_{j=0}^{J+1} w_j$ which is stable. To simplify the notation we set $P_j := P_h^{h_j} : \widetilde{V}_h^{CR} \longrightarrow W_j$, for $j = 0, \ldots, J$, and $P_{J+1} = I$, and $P_{-1} = 0$. Then for any $v \in \widetilde{V}_h^{CR}$, we set

$$v = \sum_{j=0}^{J+1} (P_j - P_{j-1}) v = \sum_{j=0}^{J} w_j, \quad \text{where} \quad w_j = (P_j - P_{j-1}) v.$$

Clearly, $w_j \in W_j$ for $j = 1, \cdots, (J+1)$ and $w_0 = P_0 v \in W_0$. Now we show that this decomposition is stable, namely that exist $C_0 > 0$ such that for all $v \in \widetilde{V}_h^{CR}$ decomposed as above,

$$a(w_0, w_0) + \sum_{j=1}^{J+1} \mathcal{R}_j(w_j, w_j) \leq C_0^2 \mathcal{A}_0(v, v) \,, \tag{5.31}$$

To estimate $\mathcal{R}_{J+1}(w_{J+1}, w_{J+1})$ we use (5.20) from Lemma 5.3, together with the approximation result (with $\widetilde{h} = h$) from Corollary 5.5 and Remark 5.6. Then by the triangle inequality, with the approximation and stability estimates of P_j from Corollary 5.5, and the fact that $\log h_j/h \simeq j$,

gives

$$\begin{aligned} a(w_0,w_0)+\sum_{j=1}^{J+1}\mathcal{R}_j(w_j,w_j) &\lesssim |P_0v|^2_{1,\kappa}+\sum_{j=1}^{J+1}h_j^{-2}\|(P_j-P_{j+1})v\|^2_{0,\kappa} \\ &\leq |P_0v|^2_{1,\kappa,\Omega}+2\sum_{j=1}^{J+1}h^{-2}\|v-P_jv\|^2_{0,\kappa} \\ &\lesssim |v|^2_{1,h,\kappa}\sum_{j=1}^{J+1}|\log 2^j| \lesssim J^2|v|^2_{1,h,\kappa}=\mathcal{A}_0(v,v)\ . \end{aligned}$$

Hence, (5.31) is shown with $C_0^2=J^2\simeq|\log h|^2$, and so $\lambda_{m_0+1}(B_{\mathrm{ML}}^{-1}A_0^{CR})\geq C_0^{-2}$.

We next give a bound on $\lambda_{max}(B_{\mathrm{ML}}^{-1}A_0^{CR})$. Let $v\in W_{J+1}=\widetilde{V}_h^{CR}$ be arbitrary and $\{w_j\}_{j=0}^{J+1}$ be any decomposition of v, namely $v=\sum_{j=0}^{J+1}w_j$, with $w_j\in W_j$. By strengthened Cauchy-Schwarz inequality, we have

$$\begin{aligned} \mathcal{A}_0(v,v) &= \mathcal{A}_0\left(\sum_{j=0}^{J+1}w_j,\sum_{j=0}^{J+1}w_j\right)\leq 3\left(\mathcal{A}_0(w_{J+1},w_{J+1})+a(w_0,w_0)+a\left(\sum_{j=1}^{J}w_j,\sum_{j=1}^{J}w_j\right)\right) \\ &= 3\left(\mathcal{A}_0(w_{J+1},w_{J+1})+a(w_0,w_0)+\sum_{i=1}^{J}\sum_{j=1}^{J}a\left(w_i,w_j\right)\right). \end{aligned}$$

By Lemma 5.9, the equivalence (5.20) and noticing that the spectral radius of the matrix $(\gamma^{|i-j|})_{J\times J}$ is uniformly bounded by $(1-\gamma)^{-1}$ we get:

$$\begin{aligned} \mathcal{A}_0(v,v) &\lesssim \mathcal{R}_{J+1}(w_{J+1},w_{J+1})+a(w_0,w_0)+\sum_{i=0}^{J-1}\sum_{j=1}^{J}\gamma^{|i-j|}\left(h_i^{-1}\|w_i\|_{0,\kappa}\right)\left(h_j^{-1}\|w_j\|_{0,\kappa}\right) \\ &\leq C_1\left(\mathcal{R}(w_{J+1},w_{J+1})+a(w_0,w_0)+\sum_{j=1}^{J}\mathcal{R}_j(w_j,w_j)\right) \end{aligned}$$

Since the decomposition was arbitrary, taking the infimum over all such decompositions together with (5.29) then gives

$$\mathcal{A}_0(v,v)\leq C_1(B_{\mathrm{ML}}^{-1}v,v),\quad \forall v\in V_h^{CR},$$

which shows that $\lambda_{\max}(B_{\mathrm{ML}}A_0^{CR})\leq C_1$, and the proof is complete. □

Remark 5.12. Similar results hold also for the multiplicative multilevel methods such as the V-cycle and these results can be easily derived from estimates comparing multiplicative and additive preconditioners given in [50, Theorem 4] or [27, Theorem 4.2].

6. Solvers for IP-1 Methods

We now introduce the different iterative methods for the solution of (2.5). As we will see, most of these methods are based on the methods constructed previously for the solution of (2.9). We begin by describing the general setting for the construction of the solvers. In all cases, we follow the ideas from [9], and we will focus on the construction of a preconditioner (iterator) denoted by B^{DG}. For simplicity we consider the following linear iteration:

Algorithm 6.1. Given initial guess u_0, for $k=0,1\ldots$ until convergence:

1. Set $e_k=B^{DG}(f-Au_k)$;
2. Update $u_{k+1}=u_k+e_k$.

Here A is the operator associated with to the bilinear form of any of the IP-1 methods.

We next discuss the construction and properties of the preconditioner (iterator) B^{DG} for symmetric and non-symmetric methods.

6.1. **Solvers for the SIPG method.** For the SIPG method ($\theta = -1$ in (2.5)), B^{DG} is used in the PCG algorithm as a preconditioner. From the spectral equivalence between $\mathcal{A}(\cdot,\cdot)$ and $\mathcal{A}_0(\cdot,\cdot)$ given in Lemma 2.2, it follows that any of the preconditioners designed for $\mathcal{A}_0(\cdot,\cdot)$ result in an efficient solver for $\mathcal{A}(\cdot,\cdot)$. In particular, due to the block diagonal form of $\mathbb{A}_0$, we focus on block-Jacobi preconditioners. We denote by R^z and $[R_0^z]$ the operator corresponding to the diagonal of $\mathcal{A}(\cdot,\cdot)$ and $\mathcal{A}_0(\cdot,\cdot)$, respectively restricted to $\mathcal{Z}_\beta$. Using the decomposition (3.3) from Proposition 3.1, we now define the following preconditioners:

$$\text{Block-Jacobi: } B_1^{DG} := [R^z]^{-1} + \widetilde{B} Q^{CR} \,, \tag{6.1}$$

$$\text{Block-Jacobi for } \mathcal{A}_0\text{: } B_0^{DG} := [R_0^z]^{-1} + \widetilde{B}_0 Q^{CR} \,,. \tag{6.2}$$

Here $Q^{CR} : V^{DG} \mapsto V_h^{CR}$ is the L_2-orthogonal projection on V_h^{CR}. Here $\widetilde{B}_0 = B_{\mathrm{ML}}$ as defined in (5.28) and $\widetilde{B}$ refers to the corresponding multilevel preconditioner for the symmetric SIPG-1 method (i.e., including the jump-jump term).

The next result is a simple consequence of the analysis given in the last two sections (Theorems 5.2 and 5.11) together with Lemma 2.2:

Theorem 6.2. *Let B^{DG} be the preconditioner defined through either* **(i)** *or* **(ii)**. *Let m_0 denote the number of floating subdomains. Then, the following estimate holds for the effective condition $\mathcal{K}_{m_0+1}(B^{DG}A)$:*

$$\mathcal{K}_{m_0+1}(B^{DG}A) \leq C|\log h|^s \lesssim |\log h|^s \,,$$

where $s = 0$ for the two level method and $s = 2$ for the multilevel method. The constant $C > 0$ above is independent of the variation in the coefficients and mesh size.

6.2. **Solvers for the non-symmetric IIPG-1 and NIPG-1 methods.** We now discuss two possible choices for the preconditioner (iterator) B^{DG} for solving the non-symmetric methods. Since the operator notation is convenient in describing these preconditioners, for a given $\mathcal{A}(\cdot,\cdot)$ (corresponding to either IIPG-1 or NIPG-1 discretization) we define the operator $A : V^{DG} \mapsto V^{DG}$ in a standard way:

$$(Av, w) := \mathcal{A}(v, w), \quad \forall v \in V^{DG}, \quad \forall w \in V^{DG}. \tag{6.3}$$

6.2.1. ***Preconditioning with the symmetric part of*** $\mathcal{A}(\cdot,\cdot)$. The first preconditioner we consider is the inverse of the symmetric part of A, and is defined as follows:

$$\begin{aligned} &B^{DG} := A_S^{-1}, \quad \text{where} \\ &(A_S v, w) := \frac{1}{2}[\mathcal{A}(v, w) + \mathcal{A}(w, v)], \quad \forall v \in V^{DG}, \quad \forall w \in V^{DG}. \end{aligned} \tag{6.4}$$

We note that from this definition and (2.15), we immediately have that A_S is symmetric and positive definite and hence defines a norm, which we will refer to as the A_S-norm and denote as $\|\cdot\|_{A_S}$. We briefly discuss and justify now why the results given in [9] can be extended to the jump coefficient problem and henceforth applied to the present situation. We mainly focus on the IIPG method since the convergence results for NIPG can be deduced by little modification of those for IIPG. We state the following result (which is an extension of [9, Theorem 5.1] to the model problem (1.1), and show uniform convergence of the linear iteration in Algorithm 6.1 with iterator B^{DG} given by (6.4)):

Theorem 6.3. *Let α^* be a fixed value of the penalty parameter for which the IIPG-0 bilinear form* (2.9) $\mathcal{A}_0(\cdot,\cdot)$ *is coercive. Let $\mathcal{A}(\cdot,\cdot)$ be the bilinear form of the IIPG-1 method* (2.5) *with penalty parameter $\alpha \geq 4\alpha^*$. Let $B^{DG} = A_S^{-1}$ be the iterator in the linear iteration 6.1 and let u_k and u_{k+1} be two consecutive iterates obtained via this algorithm. Then there exists a positive constant $\Lambda < 1$ such that*

$$\|u - u_{k+1}\|_{DG} \leq \Lambda \|u - u_k\|_{DG} \, . \tag{6.5}$$

We notice that the convergence is guaranteed under a technical (but mild) restriction on the penalty parameter, namely it should be greater than $4\alpha_*$.

A basic step in the proof of the above Theorem is to provide a uniform bound for the skew-symmetric part of $\mathcal{A}(\cdot,\cdot)$, since $([B^{DG}]^{-1}u, w) - \mathcal{A}(u,w) = ((A_S - A)u, w))$. This can be done proceeding exactly as in [9, Lemma 5.6], only notational changes are involved. In the end, the result relies on two ingredients: a *strengthened* Cauchy-Schwarz inequality that measures the angle between the two spaces in the decomposition, namely $\mathcal{Z}_\beta$ and V_h^{CR}, and a simple Lemma that uses this strengthened Cauchy-Schwarz inequality to give a further bound for the term $\mathcal{A}_0(z,v)$, $z \in \mathcal{Z}_\beta$, $v \in V_h^{CR}$. This is the step that gives rise to the technical condition on the penalty parameter. Here the key result is the fact that the functions in $\mathcal{Z}_\beta$ are orthogonal with respect to the inner product defined by $\mathcal{A}_0(\cdot,\cdot)$; or, equivalently, that the associated stiffness matrix $\mathbb{A}_0^{zz}$ is diagonal. This key result is given in Lemma 4.4 for the jump-coefficient problem. We point out that the proofs of the Cauchy-Schwarz inequality and the simple Lemma mentioned above can be carried out exactly as in [9, Lemma A & Lemma 5.1], respectively, with only small changes in the notation. As such, we omit these proofs here.

6.2.2. ***Block Jacobi preconditioner***. The second choice is simpler to implement and requires less computational work to apply. It is defined below, and is naturally referred to as a block Jacobi preconditioner.

$$\begin{aligned} &B^{DG} := D_S^{-1}, \quad \text{where} \\ &(D_S v, w) := (D_S(v^{CR} + \varphi^z), w^{CR} + \psi^z) := (A_S\varphi^z, \psi^z) + (A_S v^{CR}, w^{CR}). \end{aligned} \tag{6.6}$$

In the definition of D_S here we have set $v \in V^{DG}$, $v = v^{CR} + \varphi^z$ and $w \in V^{DG}$, $w = w^{CR} + \psi^z$, and v^{CR}, φ^z, w^{CR}, and ψ^z are the components from the decomposition (3.3). As is easily seen from the definition, D_S is the block diagonal of A_S. The corresponding matrix form of D_S is denoted by $\mathbb{D}_S$. According to the theoretical results stated and proved in Section 4 and in Section 5, we can solve a linear system with respect to D_S efficiently. We do not present a similar result here for the preconditioner B^{DG} given in (6.6). Although plausible, proving such a result does not appear to be straightforward; such estimates are the subject of current and future research. We refer to the numerical experiments Section 7 for further discussion and a comparison of the two preconditioners.

7. Numerical Experiments

We consider the model problem (1.1) in the square $\Omega = [-1,1]^2$ with coefficients:

$$\kappa(x) = \begin{cases} 1.0, & \forall x \in [-0.5,0]^2 \cup [0,0.5]^2 \\ \epsilon, & \text{elsewhere.} \end{cases}$$

In all of the following experiments, ϵ varies from 10^{-5} up to 10^5, covering a wide range of variations of the coefficients. In the experiments, we consider uniform refinement with a structured initial triangulation on level 0 with 32 elements and mesh size $h = 2^{-1}$. This initial mesh resolves the jump interface of the coefficients. Each refined triangulation is then obtained by subdividing each element of the previous level into four congruent elements. The number of degrees of freedom N_ℓ in the DG discretizations on each level satisfies $N_\ell = 4^\ell N_0$ for $\ell = 0, 1, 2, 3$

with $N_0 = 96$. For convenience a list of all tables that we have used below and the corresponding preconditioners is given in Table 7.1.

Table	Methods
7.2	two level preconditioner for the CR problem
7.3	Diagonal preconditioner for the Z part in SIPG discretization
7.4	Diagonal preconditioner for the Z part in NIPG discretization
7.5	B_1^{DG} preconditioner for SIPG discretization
7.6	A-norm of $E = (I - A_S^{-1}A)$ for IIPG with $K = 1, 2$
7.7	A-norm of $E = (I - A_S^{-1}A)$ for IIPG with $K = 4, 8$
7.8	A-norm of $E = (I - D_S^{-1}A)$ for IIPG with $K = 4, 8, 16$
7.9	A-norm of $E = (I - A_S^{-1}A)$ for NIPG with $K = 1, 2$
7.10	A-norm of $E = (I - A_S^{-1}A)$ for NIPG with $K = 4, 8, 16$
7.11	A-norm of $E = (I - D_S^{-1}A)$ for NIPG with $K = 4, 8, 16$
7.12	GMRes with preconditioner A_S^{-1} for IIPG
7.13	GMRes with preconditioner A_S^{-1} for NIPG

TABLE 7.1. List of Tables. In the list we refer to several operators: list refers to B_1^{DG} defined via (6.1); A, defined via (6.3); A_S and D_S, defined via (6.4) and (6.6).

We have used the basis (3.4)-(3.5) for our numerical tests (for all of the IP discretizations). Whenever we have used preconditioned CG or preconditioned GMRES, the iterations are stopped when the initial residual is reduced by seven orders of magnitude: namely, if r_0 is the initial residual and r^ℓ is the residual at iteration ℓ, the iteration process (PCG or preconditioned GMRES) is terminated at iteration k if $\|r^k\|_{\ell_2}/\|r^0\|_{\ell_2} < 10^{-7}$. For the non-symmetric discrete schemes (NIPG and IIPG), we have used GMRES with restart at every 10 iterations, and the maximum number of iterations is set to 30. The experiments were carried out on an IMAC (OS X) with 2.93 GHz Intel Core i7, and 8 GB 1333 MHz DDR3.

7.1. **Solver for IP(β)-0 method.** We first consider IP(β)-0 method. For this set of experiments we have set the penalty parameter $\alpha = 8$. We use algorithm 4.1 given in Section 4 to solve the linear system arising from the IP(β)-0 discretization. Due to the block structure(4.1) of $\mathbb{A}_0$ (matrix representation of A_0 in the basis (3.4)-(3.5)) we only need to numerically verify the effectiveness of the solvers for each block; $\mathbb{A}_0^{vv}$ and $\mathbb{A}_0^{zz}$. Recall that for any choice of $\theta = 0, \pm 1$, the block $\mathbb{A}_0^{vv}$ is the same (since it is the stiffness matrix of the Crouzeix-Raviart discretization (5.1)), while the block $\mathbb{A}_0^{zz}$ is different for different values of θ.

The system $\mathbb{A}_0^{vv}$ arising from the restriction of $\mathcal{A}_0(\cdot,\cdot)$ to the Crouziex-Raviart space is solved by a PCG algorithm with the two level preconditioner defined in (5.9). For the two level preconditioner we use two symmetric Gauss-Seidel steps as smoother. In Table 7.2 we report the estimated condition number $\mathcal{K}(\mathbb{B}\mathbb{A}_0^{vv})$ and the effective condition number (denoted by $\mathcal{K}_1(\mathbb{B}\mathbb{A}_0^{vv})$). Observe that the estimated condition number $\mathcal{K}(\mathbb{B}\mathbb{A}_0^{vv})$ deteriorates with respect to the magnitude of the jump in coefficient. In contrast, and as predicted by our theory, the effective condition number $\mathcal{K}(\mathbb{B}\mathbb{A}_0^{vv})$ is uniformly bounded with respect to both the mesh size and the jump of the coefficient, as predicted by Theorem 5.2. To explain the deterioration of $\mathcal{K}(\mathbb{B}\mathbb{A}_0^{vv})$ with respect to the jump in the coefficient, we have shown in Figure 7.1, the spectrum of the preconditioned system for $\epsilon = 10^{-5}$ and the mesh size $h = 2^{-5}$. Note that there is only one (very small) eigenvalue close to zero (which may be related to the fact that there are only 2 different values for the coefficients). The systems corresponding to $\mathbb{A}_0^{zz}$ are solved by a PCG al-

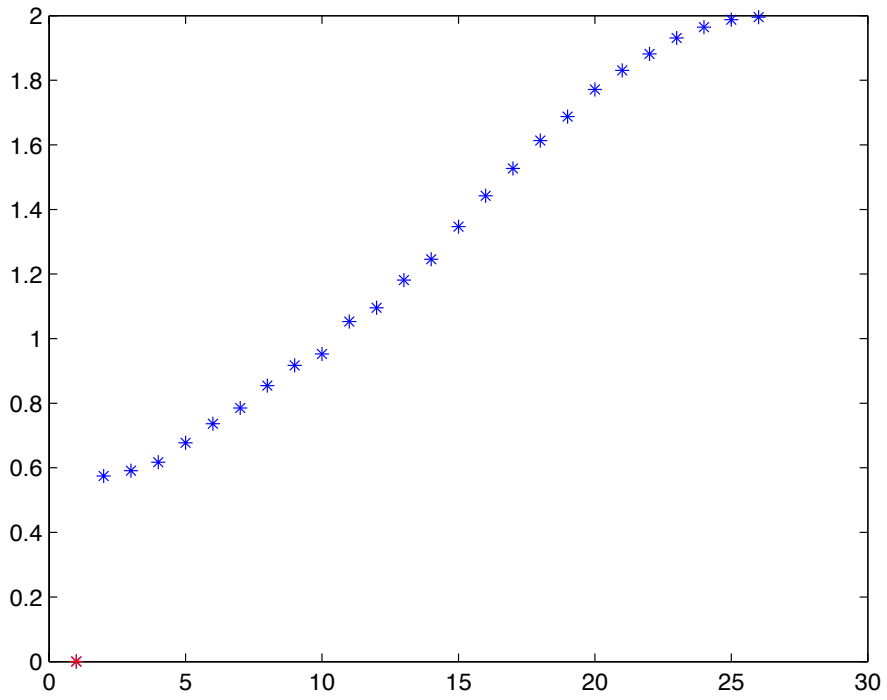


FIGURE 7.1. Eigenvalue distribution of $\mathbb{B}\mathbb{A}_0^{vv}$ for $\epsilon = 10^{-5}$ and $h = 2^{-5}$.

gorithm using its diagonal $\mathbb{D}_z$ as a preconditioner. The estimated condition numbers of $\mathbb{D}_z^{-1}\mathbb{A}_0^{zz}$ for Type-0 SIPG(β) and NIPG(β) discretizations are reported in Table 7.3 and Table 7.4, respectively. Observe that the condition numbers of $\mathbb{D}_z^{-1}\mathbb{A}_0^{zz}$ are uniformly bounded and close to 1, which confirms the result established in Lemma 4.3; i.e., that $\mathbb{A}_0^{zz}$ is spectrally equivalent to its diagonal.

7.2. **Solvers for SIPG(β)-1.** We now study the effectiveness and robustness of the proposed solvers for the SIPG-1 discretization. As before, we have set the penalty parameter $\alpha_e = 8$ in (2.5) for this set of experiments. We consider the B_1^{DG} defined via (6.1) in Section 6.1. The matrix form of this preconditioner is denoted $\mathbb{B}_1^{DG}$.

Table 7.5 gives the estimated condition numbers of $\mathcal{K}(\mathbb{B}_1^{DG}\mathbb{A})$ and the number of PCG iterations required for convergence. As can be seen from these two tables, the condition numbers of the preconditioned system deteriorate rapidly when ϵ becomes smaller. In the same table we give the effective condition numbers $\mathcal{K}_1(\mathbb{B}_1^{DG}\mathbb{A})$, and observe that the effective condition numbers are nearly uniformly bounded with respect to the coefficients and mesh size. These results verify the theory predicted by Theorem 6.2.

7.3. **Solvers for Nonsymmetric IP(β)-1 Methods.** We now present some numerical tests for the nonsymmetric IP(β)-1 methods: IIPG(β)-1 and NIPG(β)-1. We consider the linear iteration given in 6.1 with the iterator B^{DG} as defined in (6.4). As we shall see, the performance of the preconditioners depends on the value of the penalty parameter. All the tests are computed with $\alpha = K\alpha^*$, with α^* given at the beginning of §7.3. The value of α^* is close to the smallest value required for ensuring positive semi-definiteness of the symmetric part of A. In the experiments, we take $K = 1, 2, 4, 8, 16$.

Since our examples involve a fixed domain (unit square), the value of α^* can be well approximated. For the IIPG(β)-1 discretization with mesh size $h = 2^{-1}$ we take $\alpha^* = 0.9$. For all other

(a) $\epsilon < 1$

ϵ	levels	0	1	2	3
	h	2^{-1}	2^{-2}	2^{-3}	2^{-4}
10^{-5}	$\mathcal{K}(\mathbb{B}\mathbb{A}_0^{vv})$	5.49e+04 (18)	4.65e+04 (24)	3.79e+04 (26)	3.24e+04 (26)
	$\mathcal{K}_1(\mathbb{B}\mathbb{A}_0^{vv})$	4.71	4.01	3.6	3.47
10^{-4}	$\mathcal{K}(\mathbb{B}\mathbb{A}_0^{vv})$	5.49e+03 (16)	4.66e+03 (23)	3.79e+03 (23)	3.25e+03 (24)
	$\mathcal{K}_1(\mathbb{B}\mathbb{A}_0^{vv})$	4.71	4.01	3.6	3.47
10^{-3}	$\mathcal{K}(\mathbb{B}\mathbb{A}_0^{vv})$	551 (15)	469 (20)	383 (21)	328 (21)
	$\mathcal{K}_1(\mathbb{B}\mathbb{A}_0^{vv})$	4.69	4	3.6	3.47
10^{-2}	$\mathcal{K}(\mathbb{B}\mathbb{A}_0^{vv})$	57.2 (14)	49.5 (18)	41.2 (19)	36.2 (19)
	$\mathcal{K}_1(\mathbb{B}\mathbb{A}_0^{vv})$	4.52	3.95	3.57	3.44
10^{-1}	$\mathcal{K}(\mathbb{B}\mathbb{A}_0^{vv})$	7.63 (13)	7.34 (16)	6.71 (17)	6.47 (17)
	$\mathcal{K}_1(\mathbb{B}\mathbb{A}_0^{vv})$	3.48	3.54	3.32	3.27

(b) $\epsilon \geq 1$

ϵ	levels	0	1	2	3
	h	2^{-1}	2^{-2}	2^{-3}	2^{-4}
1	$\mathcal{K}(\mathbb{B}\mathbb{A}_0^{vv})$	2.58 (11)	2.77 (12)	2.79 (13)	2.94 (14)
	$\mathcal{K}_1(\mathbb{B}\mathbb{A}_0^{vv})$	2.36	2.56	2.56	2.69
10^1	$\mathcal{K}(\mathbb{B}\mathbb{A}_0^{vv})$	3.13 (10)	4.62 (15)	4.85 (16)	5.06 (16)
	$\mathcal{K}_1(\mathbb{B}\mathbb{A}_0^{vv})$	2.16	3.51	3.41	3.4
10^2	$\mathcal{K}(\mathbb{B}\mathbb{A}_0^{vv})$	3.73 (11)	6.06 (16)	7.13 (17)	8.3 (18)
	$\mathcal{K}_1(\mathbb{B}\mathbb{A}_0^{vv})$	3.36	3.8	3.68	3.66
10^3	$\mathcal{K}(\mathbb{B}\mathbb{A}_0^{vv})$	3.8 (11)	6.3 (16)	7.55 (18)	9.02 (18)
	$\mathcal{K}_1(\mathbb{B}\mathbb{A}_0^{vv})$	3.39	3.84	3.71	3.69
10^4	$\mathcal{K}(\mathbb{B}\mathbb{A}_0^{vv})$	3.8 (11)	6.32 (16)	7.59 (18)	9.09 (19)
	$\mathcal{K}_1(\mathbb{B}\mathbb{A}_0^{vv})$	3.4	3.84	3.71	3.69
10^5	$\mathcal{K}(\mathbb{B}\mathbb{A}_0^{vv})$	3.81 (11)	6.33 (16)	7.6 (18)	9.1 (20)
	$\mathcal{K}_1(\mathbb{B}\mathbb{A}_0^{vv})$	3.4	3.85	3.71	3.69

TABLE 7.2. Estimated condition numbers $\mathcal{K}(\mathbb{B}\mathbb{A}_0^{vv})$ (number of PCG iterations) and effective condition numbers $\mathcal{K}_1(\mathbb{B}\mathbb{A}_0^{vv})$ for the block $\mathbb{A}_0^{vv}$ in algorithm 4.1.

mesh sizes that we use in our numerical tests below (e.g. $h = 2^{-2}$, $h = 2^{-3}$ or $h = 2^{-4}$) we take $\alpha^* = 1.3$.

To verify the theory for Algorithm 6.1 with the symmetric part of A as an iterator, (6.4), we have computed the $\mathcal{A}$-norm of the error propagation operator. We denote this operator here with E and define it in a usual manner:

$$E = I - B^{DG}A = I - A_S^{-1}A.$$

Below, we have tabulated $\|E\|^2_{A_S}$ for several values of the parameters of interest (e.g. ϵ, mesh size, penalty parameter). This norm gives us the contraction number of the linear iteration in Algorithm 6.1. That is, the estimate in Theorem 6.3 holds with $\Lambda = \|E\|^2_{A_S}$. This norm is computed as the maximum eigenvalue of the generalized eigenvalue problem given below:

$$(I - B^{DG}A)^T A_S (I - B^{DG}A)u = \lambda A_S u,$$

where the corresponding definition of the operators A, A_S, D_S and the preconditioners B^{DG} was given earlier in (6.3), (6.4), and (6.6) respectively. For the IIPG-1 method, the A_S-norms of the error propagation operator with preconditioner (6.4) and (6.6) are given in Tables 7.6 and in Table 7.7, for the different values of K. Notice that for $K = 1$ (Table 7.6) the iteration

(a) $\epsilon < 1$

levels	h	ϵ				
		10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}
0	2^{-1}	1.73 (14)	1.73 (14)	1.73 (12)	1.73 (11)	1.73 (10)
1	2^{-2}	1.72 (15)	1.72 (14)	1.72 (13)	1.72 (12)	1.72 (10)
2	2^{-3}	1.72 (15)	1.72 (14)	1.72 (13)	1.72 (11)	1.72 (10)
3	2^{-4}	1.72 (15)	1.72 (13)	1.72 (12)	1.72 (11)	1.71 (10)

(b) $\epsilon \geq 1$

levels	h	ϵ					
		1	10^1	10^2	10^3	10^4	10^5
0	2^{-1}	1.73 (9)	1.72 (10)	1.73 (11)	1.73 (12)	1.73 (13)	1.73 (13)
1	2^{-2}	1.72 (10)	1.72 (10)	1.72 (11)	1.72 (12)	1.72 (13)	1.72 (14)
2	2^{-3}	1.71 (10)	1.7 (10)	1.72 (11)	1.71 (12)	1.72 (14)	1.72 (15)
3	2^{-4}	1.71 (10)	1.69 (10)	1.69 (11)	1.69 (12)	1.7 (14)	1.69 (16)

TABLE 7.3. Estimated condition numbers $\mathcal{K}(\mathbb{D}_z^{-1}\mathbb{A}_0^{zz})$ (number of PCG iterations) for the block $\mathbb{A}_0^{zz}$ in Type-0 SIPG(β) discretization.

(a) $\epsilon < 1$

levels	h	ϵ				
		10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}
0	2^{-1}	1.4 (12)	1.4 (12)	1.4 (10)	1.4 (10)	1.4 (8)
1	2^{-2}	1.4 (12)	1.4 (11)	1.4 (10)	1.4 (10)	1.4 (9)
2	2^{-3}	1.4 (12)	1.4 (11)	1.4 (10)	1.39 (9)	1.39 (9)
3	2^{-4}	1.4 (12)	1.39 (11)	1.39 (10)	1.39 (9)	1.38 (8)

(b) $\epsilon \geq 1$

levels	h	ϵ					
		1	10^1	10^2	10^3	10^4	10^5
0	2^{-1}	1.4 (7)	1.4 (8)	1.4 (9)	1.4 (10)	1.4 (11)	1.4 (11)
1	2^{-2}	1.39 (8)	1.38 (8)	1.38 (9)	1.38 (10)	1.4 (11)	1.4 (12)
2	2^{-3}	1.39 (8)	1.37 (8)	1.36 (9)	1.35 (10)	1.35 (12)	1.35 (13)
3	2^{-4}	1.38 (8)	1.36 (8)	1.35 (9)	1.34 (10)	1.35 (12)	1.35 (13)

TABLE 7.4. Estimated condition numbers $\mathcal{K}(\mathbb{D}_z^{-1}\mathbb{A}_0^{zz})$ (number of PCG iterations) for the block $\mathbb{A}_0^{zz}$ in Type-0 NIPG(β) discretization.

with $B^{DG} = A_S^{-1}$ is not convergent in A_S-norm. For $K \geq 4$ (as required by our theory for A_S) we have a convergent algorithm. As expected, (6.4) converges faster for larger α.

The same set of experiments is performed for the NIPG discretization. The estimates for the $\mathcal{A}$-norm of the error propagation operator corresponding to the iterator (6.4) are given in Tables 7.9 and 7.10. As can be observed from these numerical results, a larger value of K (and so of the penalty parameter) is needed than the one for the IIPG in order to produce a convergent linear iteration. From the results reported in the tables, it can be observed that (provided $K \geq 4$ for IIPG and $K \geq 16$ for NIPG) the number of iterations required for convergence does

(a) $\epsilon < 1$

ϵ	levels	0	1	2	3
	h	2^{-1}	2^{-2}	2^{-3}	2^{-4}
10^{-5}	$\mathcal{K}(\mathbb{B}_1^{DG}\mathbb{A})$	2.85e+04 (44)	3.37e+04 (44)	3.1e+04 (46)	2.85e+04 (46)
	$\mathcal{K}_1(\mathbb{B}_1^{DG}\mathbb{A})$	6.27	6.33	6.45	6.49
10^{-4}	$\mathcal{K}(\mathbb{B}_1^{DG}\mathbb{A})$	2.85e+03 (37)	3.37e+03 (38)	3.1e+03 (39)	2.86e+03 (40)
	$\mathcal{K}_1(\mathbb{B}_1^{DG}\mathbb{A})$	6.26	6.32	6.45	6.48
10^{-3}	$\mathcal{K}(\mathbb{B}_1^{DG}\mathbb{A})$	288 (33)	340 (34)	313 (34)	289 (32)
	$\mathcal{K}_1(\mathbb{B}_1^{DG}\mathbb{A})$	6.24	6.3	6.42	6.46
10^{-2}	$\mathcal{K}(\mathbb{B}_1^{DG}\mathbb{A})$	32 (27)	37 (28)	34.5 (28)	32.3 (27)
	$\mathcal{K}_1(\mathbb{B}_1^{DG}\mathbb{A})$	6.08	6.1	6.21	6.25
10^{-1}	$\mathcal{K}(\mathbb{B}_1^{DG}\mathbb{A})$	7.25 (22)	7.33 (22)	7.21 (22)	7.13 (22)
	$\mathcal{K}_1(\mathbb{B}_1^{DG}\mathbb{A})$	5.62	5.6	5.71	5.73

(b) $\epsilon \geq 1$

ϵ	levels	0	1	2	3
	h	2^{-1}	2^{-2}	2^{-3}	2^{-4}
1	$\mathcal{K}(\mathbb{B}_1^{DG}\mathbb{A})$	5.53 (19)	5.76 (20)	5.8 (20)	5.83 (20)
	$\mathcal{K}_1(\mathbb{B}_1^{DG}\mathbb{A})$	5.17	5.45	5.46	5.46
10^1	$\mathcal{K}(\mathbb{B}_1^{DG}\mathbb{A})$	6.66 (22)	7.16 (23)	7.16 (23)	7.43 (23)
	$\mathcal{K}_1(\mathbb{B}_1^{DG}\mathbb{A})$	5.91	6.2	6.25	6.27
10^2	$\mathcal{K}(\mathbb{B}_1^{DG}\mathbb{A})$	6.88 (25)	8.65 (27)	10.3 (27)	12 (28)
	$\mathcal{K}_1(\mathbb{B}_1^{DG}\mathbb{A})$	6.32	6.48	6.55	6.55
10^3	$\mathcal{K}(\mathbb{B}_1^{DG}\mathbb{A})$	6.38 (27)	8.98 (30)	11.1 (31)	13.5 (32)
	$\mathcal{K}_1(\mathbb{B}_1^{DG}\mathbb{A})$	5.51	6.53	6.59	6.59
10^4	$\mathcal{K}(\mathbb{B}_1^{DG}\mathbb{A})$	6.91 (30)	9.02 (33)	11.2 (35)	13.7 (36)
	$\mathcal{K}_1(\mathbb{B}_1^{DG}\mathbb{A})$	6.38	6.54	6.6	6.59
10^5	$\mathcal{K}(\mathbb{B}_1^{DG}\mathbb{A})$	6.91 (33)	9.02 (36)	11.3 (39)	13.8 (40)
	$\mathcal{K}_1(\mathbb{B}_1^{DG}\mathbb{A})$	6.38	6.54	6.6	6.59

TABLE 7.5. Estimated condition number $\mathcal{K}(\mathbb{B}_1^{DG}\mathbb{A})$ (number of PCG iterations) and the effective condition number $\mathcal{K}_1(\mathbb{B}_1^{DG}\mathbb{A})$.

not grow with respect to the mesh size or the coefficient jump and that $E = I - A_S^{-1}A$ is uniform contraction for such values of K.

Similar results hold for the block Jacobi preconditioner $B^{DG} = D_S^{-1}$ described in§6.2.2. From the rates presented in Table 7.8 and in Table 7.11 one may conclude that for the linear iteration with the block Jacobi preconditioner D_S, the error propagation operator $E = I - D_S^{-1}A$ is a uniform contraction in the A_S-norm. The numerical tests indicate that such estimates on the rate of convergence hold for sufficiently large values of K (same values as for the preconditioning with the symmetric part A_S) and for both IIPG and NIPG discretizations.

Once we have numerically verified that the linear iteration converges (provided $K \geq 4$ for IIPG and $K \geq 16$ for NIPG), we test the use of A_S^{-1} as a GMRES preconditioner. We have shown the number of GMRES iterations required for convergence for different values of K with the preconditioner given by (6.4) in table 7.12. The symbol **x** in Table 7.12 means that GMRES fail to converge with these parameters for > 30 iterations. It is clearly seen that preconditioned GMRES converges uniformly with respect to the mesh size and robust with respect to the jump in the coefficient, provided that $K \geq 4$ for IIPG and $K \geq 16$ for NIPG.

(a) $\|E\|^2_{A_S}$ for IIPG: $\alpha_e = \alpha^*$

levels	h	ϵ										
		10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	1	10^{1}	10^{2}	10^{3}	10^{4}	10^{5}
0	2^{-1}	2.5	2.5	2.5	2.5	2.5	2.5	2.4	2.4	2.4	2.4	2.4
1	2^{-2}	1.0	1.0	1.0	1.0	1.0	1.1	1.0	1.0	1.0	1.0	1.0
2	2^{-3}	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2	1.2

(b) $\|E\|^2_{A_S}$ for IIPG: $\alpha_e = 2\alpha^*$

levels	h	ϵ										
		10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	1	10^{1}	10^{2}	10^{3}	10^{4}	10^{5}
0	2^{-1}	0.53	0.53	0.53	0.53	0.53	0.53	0.51	0.52	0.52	0.52	0.52
1	2^{-2}	0.33	0.33	0.33	0.33	0.34	0.34	0.33	0.34	0.34	0.34	0.34
2	2^{-3}	0.37	0.37	0.37	0.37	0.37	0.37	0.36	0.37	0.37	0.37	0.37
3	2^{-4}	0.39	0.39	0.39	0.39	0.39	0.39	0.38	0.39	0.39	0.39	0.39

TABLE 7.6. Norm of the error propagator $E = (I - A_S^{-1}A)$ for A corresponding to IIPG discretization, with: (a) $\alpha_e = \alpha^*$ and (b) $\alpha_e = 2\alpha^*$.

(a) $\|E\|^2_{A_S}$ for IIPG: $\alpha_e = 4\alpha^*$

levels	h	ϵ										
		10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	1	10^{1}	10^{2}	10^{3}	10^{4}	10^{5}
0	2^{-1}	0.20	0.20	0.20	0.20	0.20	0.20	0.19	0.19	0.19	0.19	0.19
1	2^{-2}	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14
2	2^{-3}	0.16	0.16	0.16	0.16	0.16	0.15	0.15	0.16	0.16	0.16	0.16
3	2^{-4}	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16

(b) $\|E\|^2_{A_S}$ for IIPG: $\alpha_e = 8\alpha^*$

levels	h	ϵ										
		10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	1	10^{1}	10^{2}	10^{3}	10^{4}	10^{5}
0	2^{-1}	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
1	2^{-2}	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07
2	2^{-3}	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07
3	2^{-4}	0.08	0.08	0.08	0.08	0.08	0.07	0.07	0.07	0.07	0.07	0.07

TABLE 7.7. Norm of the error propagator $E = (I - A_S^{-1}A)$ for A corresponding to IIPG discretization, with: (a) $\alpha_e = 4\alpha^*$ and (b) $\alpha_e = 8\alpha^*$.

We conduct the same set of experiments for NIPG(β) discretization. The number of GMRES iterations required for convergence for different values of K with the preconditioner (6.4) is shown in Table 7.13. Clearly with such a preconditioner the GMRES method is uniformly convergent with respect to the problem parameters for $K \geq 16$.

(a) $\|E\|^2_{A_S}$ for IIPG: $\alpha_e = 4\alpha^*$

		ϵ										
levels	h	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	1	10^{1}	10^{2}	10^{3}	10^{4}	10^{5}
0	2^{-1}	0.81	0.81	0.81	0.81	0.80	0.80	0.80	0.80	0.80	0.80	0.80
1	2^{-2}	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67
2	2^{-3}	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68
3	2^{-4}	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68	0.68

(b) $\|E\|^2_{A_S}$ for IIPG: $\alpha_e = 8\alpha^*$

		ϵ										
levels	h	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	1	10^{1}	10^{2}	10^{3}	10^{4}	10^{5}
0	2^{-1}	0.59	0.59	0.59	0.59	0.59	0.58	0.58	0.58	0.59	0.59	0.59
1	2^{-2}	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55
2	2^{-3}	0.56	0.56	0.56	0.56	0.56	0.56	0.56	0.56	0.56	0.56	0.56
3	2^{-4}	0.56	0.56	0.56	0.56	0.56	0.56	0.56	0.56	0.56	0.56	0.56

(c) $\|E\|^2_{A_S}$ for IIPG: $\alpha_e = 16\alpha^*$

		ϵ										
levels	h	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	1	10^{1}	10^{2}	10^{3}	10^{4}	10^{5}
0	2^{-1}	0.52	0.52	0.52	0.52	0.51	0.50	0.50	0.50	0.50	0.50	0.50
1	2^{-2}	0.51	0.51	0.51	0.51	0.51	0.51	0.50	0.50	0.50	0.50	0.50
2	2^{-3}	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51
3	2^{-4}	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.51

TABLE 7.8. Norm of the error propagator $E = (I - D_S^{-1}A)$ for A corresponding to IIPG discretization, with: (a) $\alpha_e = 4\alpha^*$, (b) $\alpha_e = 8\alpha^*$ and (c) $\alpha_e = 16\alpha^*$. .

APPENDIX A. CONSTRUCTION OF THE TRANSFER OPERATOR $P_h^{\widetilde{h}}$.

We now give the definition of an operator $P_h^{\widetilde{h}}$ that will be shown to satisfy the required approximation and stability properties (5.23)-(5.24). We start by introducing some notation required for our construction. For a given conforming triangulation $\mathcal{T}_h$, we denote by $\mathcal{N}(\mathcal{T}_h)$ the set of vertices of the partition $\mathcal{T}_h$, and by $\mathcal{C}(\mathcal{T}_h)$ the set of barycenters of the elements $T \in \mathcal{T}_h$. We still denote by $\mathcal{E}_h$ the set of edges/faces of $\mathcal{T}_h$. For each vertex $p \in \mathcal{N}$, let $\omega_p := \bigcup_{T \ni p} T$ and $\omega_T = \bigcup_{p \in T} \omega_p$ for each $T \in \mathcal{T}_h$. Similarly, on the interface $\mathcal{O}_p$ and $\mathcal{O}_e$ denote, respectively, the local patches associated with the vertex $p \in \mathcal{N}$ and the edge/face $e \in \mathcal{E}_h$ on the interface. We now start building the operator $P_h^{\widetilde{h}}$; it is constructed in several steps as the composition of a particular *inclusion operator* and the Scott-Zhang quasi-interpolation operator. The first basic idea is to embed V_h^{CR} into a higher order conforming finite element space on the same mesh $\mathcal{T}_h$. Following [17], we consider the space of piecewise quadratic polynomials on $\mathcal{T}_h$, which we denote by $V_h^{\text{conf},2}$. To be able to use the results in [17] for the jump-coefficients problem, we consider this inclusion at the subdomain level. Let $E_i : V_h^{CR}(\Omega_i) :\to V_h^{\text{conf},2}(\Omega_i)$ be the inclusion

(a) $\|E\|^2_{A_S}$ for NIPG: $\alpha_e = \alpha^*$

levels	h	ϵ										
		10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	1	10^{1}	10^{2}	10^{3}	10^{4}	10^{5}
0	2^{-2}	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7	2.7
1	2^{-3}	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
2	2^{-4}	2.2	2.1	2.1	2.2	2.2	2.1	2.1	2.1	2.2	2.2	2.2

(b) $\|E\|^2_{A_S}$ for NIPG: $\alpha_e = 2\alpha^*$

levels	h	ϵ										
		10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	1	10^{1}	10^{2}	10^{3}	10^{4}	10^{5}
0	2^{-1}	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3
1	2^{-2}	0.98	0.98	0.98	0.98	0.99	0.99	0.97	0.98	0.99	0.99	0.99
2	2^{-3}	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1
3	2^{-4}	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1

TABLE 7.9. Norm of the error propagator $E = (I - A_S^{-1}A)$ for A corresponding to NIPG discretization, with: (a) $\alpha_e = \alpha^*$ and (b) $\alpha_e = 2\alpha^*$.

operator defined on each subdomain Ω_i by:

$$E_i[v(p)] = \frac{1}{|\omega_p|}\sum_{T\in\omega_p} v_T(p) \qquad \forall p \in \mathcal{N}\left(\mathcal{T}_h(\Omega_i)\right) \cup \mathcal{C}\left(\mathcal{T}_h(\Omega_i)\right) \ . \tag{A.1}$$

The above definition applies for all interior points $p \in \mathcal{N}\left(\mathcal{T}_h(\Omega_i)\right) \cup \mathcal{C}\left(\mathcal{T}_h(\Omega_i)\right)$. We introduce the set $\Xi_p^i := \{T \in \mathcal{T}_h \,:\, p \in \partial T \text{ and } \text{meas}_{d-1}(\partial T \cap \partial\Omega_i) > 0\}$ and we note that the cross-points do not belong to this set. At the boundary vertices $p \in \partial\Omega_i$, we define E_i in the following way:

$$(E_i v)(p) = \frac{1}{|\Xi_p^i|}\sum_{T\in\Xi_p^i} v_T(p) \quad \forall p \in \mathcal{N}\left(\mathcal{T}_h(\partial\Omega_i)\right) \cup \mathcal{C}\left(\mathcal{T}_h(\partial\Omega_i)\right).$$

Observe that, with this construction, we have that $\eta_i = E_i v \in V_h^{\text{conf},2}(\Omega_i)$ is conforming at each subdomain. However, the global function $\eta|_{\Omega_i} := \eta_i$ will generally be multi-valued at the cross points. The modification of the definition of E_i at boundary points is done to guarantee that the global function η will be at least conforming every non-cross point of the partition and so, in particular, in the interior of the boundary of each subdomain. The next result is taken from [17], where the author studies uniform preconditioners and solvers for the CR discretization of the Poisson problem. It guarantees the stability and approximation properties of the *inclusion* operator E_i. The proof can be found in [17]:

Lemma A.1. *There exists a linear operator $E_i : V_h^{CR}(\Omega_i) :\to V_h^{\text{conf},2}(\Omega_i)$ defined on each subdomain Ω_i such that for any $v \in V_h^{CR}$*

(i) $|E_i v|_{1,\Omega_i} \simeq |v|_{1,h,\Omega_i}$
(ii) $\|v - E_i v\|_{0,\Omega_i} \lesssim h\,|v|_{1,h,\Omega_i}$

We note that $V_{\tilde{h}}^{\text{conf}} \subset V_h^{\text{conf},2}$ and therefore a standard restriction operator would allow us to represent the constructed η as a function in $V_{\tilde{h}}^{\text{conf}}$ (except at cross-points). However, to guarantee that the constructed operator satisfies (5.23)-(5.24), we use the *Scott-Zhang quasi-interpolation operator*, again at the subdomain level and also at the interior of the interfaces, say $\Gamma = \partial\Omega_i \cap \partial\Omega_j \;\; i \neq j$, between any pair of subdomains. We denote by $\mathcal{Q}_i : H^1(\Omega_i) \to V_{\tilde{h}}^{\text{conf}}(\Omega_i)$

(a) $\|E\|^2_{A_S}$ for NIPG: $\alpha_e = 4\alpha^*$

		ϵ										
levels	h	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	1	10^{1}	10^{2}	10^{3}	10^{4}	10^{5}
0	2^{-1}	0.66	0.66	0.66	0.66	0.66	0.66	0.64	0.64	0.64	0.64	0.64
1	2^{-2}	0.49	0.49	0.49	0.49	0.49	0.49	0.48	0.49	0.49	0.49	0.49
2	2^{-3}	0.54	0.54	0.54	0.54	0.54	0.53	0.52	0.53	0.54	0.54	0.54
3	2^{-4}	0.56	0.56	0.56	0.56	0.56	0.55	0.55	0.56	0.56	0.56	0.56

(b) $\|E\|^2_{A_S}$ for NIPG: $\alpha_e = 8\alpha^*$

		ϵ										
levels	h	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	1	10^{1}	10^{2}	10^{3}	10^{4}	10^{5}
0	2^{-1}	0.32	0.32	0.32	0.32	0.32	0.32	0.31	0.31	0.31	0.31	0.31
1	2^{-2}	0.24	0.24	0.24	0.24	0.24	0.24	0.24	0.24	0.24	0.24	0.24
2	2^{-3}	0.27	0.27	0.27	0.27	0.27	0.26	0.26	0.27	0.27	0.27	0.27
3	2^{-4}	0.28	0.28	0.28	0.28	0.28	0.28	0.27	0.28	0.28	0.28	0.28

(c) $\|E\|^2_{A_S}$ for NIPG: $\alpha_e = 2\alpha^*$

		ϵ										
levels	h	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	1	10^{1}	10^{2}	10^{3}	10^{4}	10^{5}
0	2^{-1}	0.16	0.16	0.16	0.16	0.16	0.16	0.15	0.15	0.15	0.15	0.15
1	2^{-2}	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12
2	2^{-3}	0.14	0.14	0.14	0.14	0.14	0.13	0.13	0.13	0.13	0.13	0.13
3	2^{-4}	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14

TABLE 7.10. Norm of the error propagator $E = (I - A_S^{-1}A)$ for A corresponding to NIPG discretization, with: (a) $\alpha_e = 4\alpha^*$; (b) $\alpha_e = 8\alpha^*$; and (c) $\alpha_e = 16\alpha^*$.

the Scott-Zhang quasi-interpolation operator in Ω_i and we let $\mathcal{Q}_\Gamma : L^2(\Gamma) \to V^{\text{conf}}_{\widetilde{h}}(\Gamma)$ be the corresponding Scott-Zhang operator on $\Gamma \subset \partial\Omega_i$. We now recall the definition and the main properties of these interpolators. For any $p \in \mathcal{N}\left(\mathcal{T}_{\widetilde{h}}(\Omega_i)\right)$, choose some[1] $T \subset \omega_p$. Let $\{\lambda_{T,i} : i = 1, \cdots, d+1\}$ be the barycentric coordinates of T and let denote by $\{\theta_{T,i} : i = 1, \cdots, d+1\}$ its L^2-dual basis, i.e., $(\lambda_{T,i}, \theta_{T,j})_T = \delta_{i,j}$. Let $\{\phi_p\}_{p\in\mathcal{N}(\mathcal{T}_{\widetilde{h}}(\Omega_i))}$ denote the set of nodal basis of $V^{\text{conf}}_{\widetilde{h}}(\Omega_i)$. Then, the Scott-Zhang quasi-interpolation operator is defined by

$$\mathcal{Q}_i\eta = \sum_{p\in\mathcal{N}\left(\mathcal{T}_{\widetilde{h}}(\Omega_i)\right)} \left(\int_T \theta_p\eta\right)\phi_p \,,$$

The operator $\mathcal{Q}_\Gamma$ is defined similarly. Both operators enjoy the following approximation and stability properties:

Lemma A.2. *For any $\eta \in H^1(\Omega_i)$, the quasi-interpolation operators $\mathcal{Q}_i : H^1(\Omega_i) \to V^{\text{conf}}_{\widetilde{h}}(\Omega_i)$ and $\mathcal{Q}_\Gamma : L^2(\Gamma) \to V^{\text{conf}}_{\widetilde{h}}(\Gamma)$ with $\Gamma \subset \partial\Omega_i$, satisfy the following properties:*

$$\|\mathcal{Q}_i\eta\|_{0,T} \lesssim \|\eta\|_{0,\omega_T}, \quad \|\mathcal{Q}_i\eta\|_{1,T} \lesssim \|\eta\|_{1,\omega_T}\,, \qquad \|\mathcal{Q}_\Gamma\eta\|_{0,F} \lesssim \|\eta\|_{0,\mathcal{O}_F}\,. \tag{A.2}$$

$$\|\widetilde{h}^{-1}(I - \mathcal{Q}_i\eta)\|_{0,T} \lesssim \|\eta\|_{1,\omega_T}\,, \qquad \|\widetilde{h}^{-1}(I - \mathcal{Q}_\Gamma\eta)\|_{0,F} \lesssim \|\eta\|_{1,\mathcal{O}_F}\,. \tag{A.3}$$

[1]Note that the choice of T is not unique

(a) $\|E\|^2_{A_S}$ for NIPG: $\alpha_e = 4\alpha^*$

		ϵ										
levels	h	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	1	10^{1}	10^{2}	10^{3}	10^{4}	10^{5}
0	2^{-1}	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.6	1.6
1	2^{-2}	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3	1.3
2	2^{-3}	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
3	2^{-4}	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4

(b) $\|E\|^2_{A_S}$ for NIPG: $\alpha_e = 8\alpha^*$

		ϵ										
levels	h	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	1	10^{1}	10^{2}	10^{3}	10^{4}	10^{5}
0	2^{-1}	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1	2^{-2}	0.83	0.83	0.83	0.83	0.83	0.83	0.82	0.83	0.83	0.83	0.83
2	2^{-3}	0.86	0.86	0.86	0.86	0.86	0.86	0.85	0.86	0.86	0.86	0.86
3	2^{-4}	0.88	0.88	0.88	0.88	0.88	0.88	0.87	0.88	0.88	0.88	0.88

(c) $\|E\|^2_{A_S}$ for NIPG: $\alpha_e = 16\alpha^*$

		ϵ										
levels	h	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	1	10^{1}	10^{2}	10^{3}	10^{4}	10^{5}
0	2^{-1}	0.68	0.68	0.68	0.68	0.68	0.67	0.68	0.68	0.68	0.68	0.68
1	2^{-2}	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60	0.60
2	2^{-3}	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61
3	2^{-4}	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61	0.61

TABLE 7.11. Norm of the error propagator $E = (I - D_S^{-1}A)$ for A corresponding to NIPG discretization, with: (a) $\alpha_e = 4\alpha^*$, (b) $\alpha_e = 8\alpha^*$ and (c) $\alpha_e = 16\alpha^*$. .

Furthermore, both operators are linear preserving; i.e. if $\eta \in V_{\widetilde{h}}^{\text{conf}}(\omega_T)$, *then* $\mathcal{Q}_i\eta(x) \equiv \eta(x)$, *and similarly if* $\eta \in V_{\widetilde{h}}^{\text{conf}}(\mathcal{O}_e)$, *then* $\mathcal{Q}_\Gamma\eta(x) \equiv \eta(x)$.

We refer to [60, 66] for a proof of the above result.

By using $\mathcal{Q}_i$ and $\mathcal{Q}_\Gamma$, we can now define our local interpolation operator $P_h^{\widetilde{h}}$, for which the properties (5.23)-(5.24) can be shown. The definition is inspired on the ideas from [15] for the weighted L^2-projection. To define such an operator, we treat the interior of each subdomain and the interfaces differently. We define the interpolation operator $P_h^{\widetilde{h}} : V_h^{CR} \to V_{\widetilde{h}}^{\text{conf}}$ as follows:

$$P_h^{\widetilde{h}} v = \begin{cases} \mathcal{Q}_i E_i v, & \text{at vertices inside the subdomain } \Omega_i \\ \mathcal{Q}_\Gamma E_i v, & \text{at vertices inside each face } \Gamma \text{ of the subdomain } \Omega_i \\ 0, & \text{at vertices elsewhere (cross-points).} \end{cases} \tag{A.4}$$

The next result guarantees that the operator $P_h^{\widetilde{h}}$ defined above does satisfy the approximation and stability properties (5.23)-(5.24):

Lemma A.3. *For any* $v \in V_h^{CR}$, *the operator* $P_h^{\widetilde{h}} : V_h^{CR} \to V_{\widetilde{h}}^{\text{conf}}$ *satisfies*

$$\|(I - P_h^{\widetilde{h}})v\|_{0,\kappa} \lesssim \widetilde{h}|\log 2\widetilde{h}/h|^{1/2}\|v\|_{1,h,\kappa}, \tag{A.5}$$

$$|P_h^{\widetilde{h}} v|_{1,\kappa} \lesssim |\log 2\widetilde{h}/h|^{1/2}\|v\|_{1,h,\kappa}\,. \tag{A.6}$$

(a) Preconditioned GMRES for IIPG: $\alpha_e = \alpha^*$

		ϵ										
levels	h	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	1	10^{1}	10^{2}	10^{3}	10^{4}	10^{5}
0	2^{-1}	x	x	6	5	4	3	3	2	2	2	1
1	2^{-2}	x	6	5	4	3	2	2	2	2	2	2
2	2^{-3}	x	x	5	4	3	3	2	2	2	2	2
3	2^{-4}	x	x	5	4	3	3	2	2	2	2	2

(b) Preconditioned GMRES for IIPG: $\alpha_e = 2\alpha^*$

		ϵ										
levels	h	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	1	10^{1}	10^{2}	10^{3}	10^{4}	10^{5}
0	2^{-1}	x	4	3	3	2	2	2	1	1	1	1
1	2^{-2}	x	4	3	2	2	2	2	1	1	1	1
2	2^{-3}	x	4	3	3	2	2	2	2	2	1	1
3	2^{-4}	x	4	3	3	2	2	2	2	2	2	2

(c) Preconditioned GMRES for IIPG: $\alpha_e = 4\alpha^*$

		ϵ										
levels	h	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	1	10^{1}	10^{2}	10^{3}	10^{4}	10^{5}
0	2^{-1}	4	3	3	2	2	1	1	1	1	1	1
1	2^{-2}	30	3	2	2	2	1	1	1	1	1	1
2	2^{-3}	16	3	2	2	2	2	1	1	1	1	1
3	2^{-4}	11	3	2	2	2	2	1	1	1	1	1

TABLE 7.12. Number of GMRES iterations with the preconditioner A_S^{-1} for IIPG, with: (a) $\alpha_e = \alpha^*$; (b) $\alpha_e = 2\alpha^*$; and (c) $\alpha_e = 4\alpha^*$.

The proof follows the ideas from [15, Lemma 4.6].

Proof. We start showing the approximation estimate. Using standard triangle inequality, the stability and approximation properties of E_i given in Lemma A.1, together with the approximation result (A.3) of the $\mathcal{Q}_i$ from Lemma A.2, we have

$$\begin{aligned}
\|v - P_h^{\widetilde{h}} v\|_{0,\Omega_i} &\leq \|v - \mathcal{Q}_i E_i v\|_{0,\Omega_i} + \|\mathcal{Q}_i E_i v - P_h^{\widetilde{h}} v\|_{0,\Omega_i} \\
&\lesssim \|v - E_i v\|_{0,\Omega_i} + \|(I - \mathcal{Q}_i) E_i v\|_{0,\Omega_i} + \|\mathcal{Q}_i E_i v - P_h^{\widetilde{h}} v\|_{0,\Omega_i} \\
&\lesssim h|v|_{1,h,\Omega_i} + \widetilde{h}\|E_i v\|_{1,\Omega_i} + \|\mathcal{Q}_i E_i v - P_h^{\widetilde{h}} v\|_{0,\Omega_i} , \\
&\lesssim h|v|_{1,h,\Omega_i} + \widetilde{h}\|v\|_{1,h,\Omega_i} + \|\mathcal{Q}_i E_i v - P_h^{\widetilde{h}} v\|_{0,\Omega_i} , \qquad \text{(A.7)}
\end{aligned}$$

Hence, to conclude we only need to estimate $\|\mathcal{Q}_i E_i v - P_h^{\widetilde{h}} v\|_{0,\Omega_i}$. To simplify the notation, throughout the proof we set $\chi = P_{\widetilde{h}}^{\kappa} v \in V_{\widetilde{h}}^{\text{conf}}$ as defined in (A.4), and denote $\chi_i := \mathcal{Q}_i E_i v$.

(a) Preconditioned GMRES for NIPG: $\alpha_e = 8\alpha^*$

		ϵ										
levels	h	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	1	10^{1}	10^{2}	10^{3}	10^{4}	10^{5}
0	2^{-1}	4	3	3	2	2	2	1	1	1	1	1
1	2^{-2}	x	3	3	2	2	2	1	1	1	1	1
2	2^{-3}	x	3	3	2	2	2	2	1	1	1	1
3	2^{-4}	x	3	3	2	2	2	2	2	1	1	1

(b) Preconditioned GMRES for NIPG: $\alpha_e = 16\alpha^*$

		ϵ										
levels	h	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}	1	10^{1}	10^{2}	10^{3}	10^{4}	10^{5}
0	2^{-1}	3	3	2	2	2	1	1	1	1	1	1
1	2^{-2}	5	3	2	2	2	1	1	1	1	1	1
2	2^{-3}	3	3	2	2	2	1	1	1	1	1	1
3	2^{-4}	4	3	2	2	2	1	1	1	1	1	1

TABLE 7.13. Number of GMRES iterations with the preconditioner A_S^{-1} for NIPG, with: (a) $\alpha_e = 8\alpha^*$ and (b) $\alpha_e = 16\alpha^*$.

Then by using discrete L^2 norm, we have

$$\begin{aligned}
\|\mathcal{Q}_i E_i v - P_h^{\widetilde{h}} v\|_{0,\Omega_i} &= \|\chi - \chi_i\|^2_{0,\Omega_i} \qquad \text{(A.8)}\\
&\lesssim \sum_{p\in\partial\Omega_i\cap\mathcal{N}_{\widetilde{h}}(\Omega_i} \widetilde{h}^d(\chi-\chi_i)^2(p)\\
&\lesssim \sum_{\Gamma\subset\partial\Omega_i}\sum_{p\in\Gamma} \widetilde{h}^d(\chi-\chi_i)^2(p)\\
&\lesssim \sum_{\Gamma\subset\partial\Omega_i}\left(\sum_{p\in\Gamma}\widetilde{h}^d\left(\chi_i - \mathcal{Q}_\Gamma E_i v\right)^2(p) + \sum_{p\in\partial\Gamma}\widetilde{h}^d\chi_i^2(p)\right)\\
&\lesssim \sum_{\Gamma\in\partial\Omega_i}\left(\sum_{e\subset\Gamma}\widetilde{h}\|(\chi_i - \mathcal{Q}_\Gamma E_i v)\|^2_{0,e} + \widetilde{h}^2\|\chi_i\|^2_{0,\partial\Gamma}\right). \qquad \text{(A.9)}
\end{aligned}$$

We need to bound the two terms appearing in the last expression. Taking into account that $\mathcal{Q}_\Gamma$ is linear preserving (cf. Lemma A.2) and also using its L^2-stability property (A.2), the standard trace inequality, and the local approximation property of the Scott-Zhang interpolation $\mathcal{Q}_i$ (A.3), we find for the first term

$$\begin{aligned}
\widetilde{h}\|\chi_i - \mathcal{Q}_\Gamma E_i v\|^2_{0,e} &= \widetilde{h}\|\mathcal{Q}_\Gamma\chi_i - \mathcal{Q}_\Gamma E_i v\|^2_{0,e} \lesssim \widetilde{h}\|\chi_i - E_i v\|^2_{0,\mathcal{O}_e}\\
&\lesssim \widetilde{h}\left(\operatorname{diam}(\omega_{\mathrm{p}})\right)^{-1}\|E_i v - \mathcal{Q}_i E_i v\|^2_{0,\omega_p} + \widetilde{h}\operatorname{diam}(\omega_{\mathrm{p}})|E_i v - \mathcal{Q}_i E_i v|^2_{1,\omega_p}\\
&\lesssim \widetilde{h}\operatorname{diam}(\omega_{\mathrm{p}})\|E_i v\|^2_{1,\omega_p},
\end{aligned}$$

where $\mathcal{O}_e \subset \partial\Omega_i$ is the local patch associated with $e \in \mathcal{T}_{\widetilde{h}}|_{\partial\Omega_i}$ on the interface, and $\omega_p \subset \Omega_i$ is the patch associated with the vertex $p \in e$. Note that $\operatorname{diam}(\omega_{\mathrm{p}}) = 2\widetilde{h}$.

Summing up the above inequality for all edges/faces, we obtain that

$$\sum_{\Gamma\subset\partial\Omega_i}\sum_{e\subset\Gamma}\widetilde{h}\|\chi_i - \mathcal{Q}_\Gamma E_i v\|^2_{0,e} \lesssim \widetilde{h}\operatorname{diam}(\omega_{\mathrm{p}})\|E_i v\|^2_{1,\Omega_i} \lesssim \widetilde{h}\operatorname{diam}(\omega_{\mathrm{p}})\|v\|^2_{1,h,\Omega_i}. \qquad \text{(A.10)}$$

To bound the second term in (A.9) we have to distinguish between the $2D$ and $3D$ cases. In the $2D$ case, Γ is one-dimensional, (it is a set of "edges") and so $\partial\Gamma$ reduces to its two endpoints, say $\{p, q\}$. Hence,

$$\|\chi_i\|^2_{0,\partial\Gamma} = (|\chi_i(p)|^2 + |\chi_i(q)|^2) \leq \|\chi_i\|^2_{0,\infty,\omega_p} + \|\chi_i\|^2_{0,\infty,\omega_q} , \qquad \partial\Gamma = \{p, q\} ,$$

To bound each of the above two terms on the right side, we use the two-dimensional discrete Sobolev inequality [15, Lemma 2.3];

$$\|\chi_i\|_{0,\infty,\omega_p} \leq C \left(\log \frac{\mathrm{diam}(\omega_\mathrm{p})}{h}\right)^{1/2} \|\chi_i\|_{1,\omega_p} , \tag{A.11}$$

and so summing over all $\Gamma \in \partial\Omega_i$ the resulting estimate, and using the stability (A.2) from Lemma A.2 of $\mathcal{Q}_i$ together with the stability and approximation properties of E_i given in Lemma A.1, we finally get

$$\begin{aligned}\sum_{\Gamma\in\partial\Omega_i} \|\chi_i\|^2_{0,\partial\Gamma} &\lesssim \sum_{\Gamma\in\partial\Omega_i}\sum_{p\subset\partial\Gamma} \log\left(\frac{\mathrm{diam}(\omega_\mathrm{p})}{h}\right)\|\chi_i\|^2_{1,\omega_p} \lesssim \log\left(\frac{2\widetilde{h}}{h}\right)\|\chi_i\|^2_{1,\Omega_i} \\ &= \log\left(\frac{2\widetilde{h}}{h}\right)\|\mathcal{Q}_i E_i v\|^2_{1,\Omega_i} \lesssim \log\left(\frac{2\widetilde{h}}{h}\right)\|v\|^2_{1,\Omega_i} .\end{aligned} \tag{A.12}$$

Next, we give the corresponding estimate for the 3-dimensional case. In $3D$, any $\Gamma \subset \partial\Omega_i$ is a set of faces and edges. Notice however, that it is enough to consider the case when Γ is two-dimensional (face), since the other case reduces to the estimate already done for $2D$. Γ being a "face", $\partial\Gamma$ reduces to a set of edges, $\{e \ : \ e \subset \partial\Gamma\}$. Hence, defining the set $\omega_e := \{T \in \mathcal{T}_{\widetilde{h}} \ : \ T \cap e \neq \emptyset \ e \subset \partial\Gamma\}$ and using now the discrete Sobolev inequality [15, Lemma 2.4] (instead of (A.11)), we find

$$\|\chi_i\|^2_{0,\partial\Gamma} = \sum_{e\subset\partial\Gamma} \|\chi_i\|^2_{0,e} \lesssim \sum_{e\subset\partial\Gamma} \log\left(\frac{\mathrm{diam}(\omega_\mathrm{e})}{h}\right)\|\chi_i\|^2_{1,\omega_e} .$$

Summing the above estimate over all $\Gamma \in \partial\Omega_i$ and using, as before, the stability (A.2) from Lemma A.2 of $\mathcal{Q}_i$ together with the stability and approximation properties of E_i given in Lemma A.1, we find,

$$\begin{aligned}\sum_{\Gamma\in\partial\Omega_i} \|\chi_i\|^2_{0,\partial\Gamma} &\lesssim \sum_{\Gamma\in\partial\Omega_i}\sum_{e\subset\partial\Gamma} \log\left(\frac{\mathrm{diam}(\omega_\mathrm{e})}{h}\right)\|\chi_i\|^2_{1,\omega_e} \lesssim \log\left(\frac{2\widetilde{h}}{h}\right)\sum_{\Gamma\in\partial\Omega_i}\sum_{e\subset\partial\Gamma}\|\chi_i\|^2_{1,\omega_e} \\ &\lesssim \log\left(\frac{2\widetilde{h}}{h}\right)\|\chi_i\|^2_{1,\Omega_i} \lesssim \log\left(\frac{2\widetilde{h}}{h}\right)\|v\|^2_{1,\Omega_i} .\end{aligned}$$

Now, substituting into (A.9) the above estimate (or correspondingly (A.12)) together with (A.10) we finally get

$$\|\mathcal{Q}_i E_i v - P_h^{\widetilde{h}} v\|^2_{0,\Omega_i} = \|\chi - \chi_i\|^2_{0,\Omega_i} \leq C2C\widetilde{h}^2\|v\|^2_{1,h,\Omega_i} + C\widetilde{h}^2 \log\left(\frac{2\widetilde{h}}{h}\right)\|v\|^2_{1,\Omega_i} .$$

Inserting this estimate in (A.7), the proof of the approximation property is concluded by using now the definition of the H^1-weighted norm.

Finally we show the stability of $P_h^{\widetilde{h}}$ (A.6). From the definition of the norm

$$\|P_h^{\widetilde{h}} v\|^2_{1,\kappa,\Omega} = \sum_{\Omega_i\subset\Omega} \kappa\|P_h^{\widetilde{h}} v\|^2_{1,\Omega_i} , \quad \|P_h^{\widetilde{h}} v\|^2_{1,\Omega_i} = \sum_{T\subset\Omega_i,\bigcup T=\Omega_i} \|P_h^{\widetilde{h}} v\|^2_{1,T} .$$

Note that $P_h^{\widetilde{h}} v \in V_{\widetilde{h}}^{\text{conf}}$ and $v \in V_h^{CR}$. To deal with possibly different mesh sizes we consider the local L^2-projection $\mathcal{P}_T : L^2(T) \longrightarrow \mathbb{P}^1(T)$ for any $T \in \mathcal{T}_{\widetilde{h}}$. For $\widetilde{h} > h$, such an element is the union of other subelements in the partition $\mathcal{T}_h$. Then, adding and subtracting $\mathcal{P}_T v$, triangle inequality together with inverse inequality and the approximation property (A.5), gives

$$
\begin{aligned}
|P_h^{\widetilde{h}} v|_{1,T} &\leq |P_h^{\widetilde{h}} v - \mathcal{P}_T v|_{1,T} + |\mathcal{P}_T v|_{1,T} \leq C(\widetilde{h})^{-1} \|P_h^{\widetilde{h}} v - \mathcal{P}_T v\|_{0,T} + |\mathcal{P}_T v|_{1,T} \\
&\leq C(\widetilde{h})^{-1} \left(\|P_h^{\widetilde{h}} v - v\|_{0,T} + \|v - \mathcal{P}_T v\|_{0,T} \right) + C|v|_{1,T} \leq C(\widetilde{h})^{-1} \|P_h^{\widetilde{h}} v - v\|_{0,T} + C\|v\|_{1,T} \, .
\end{aligned}
$$

The Stability now follows immediately, by summing over all elements $T \subset \Omega_i$, using the definition of the weighted H^1-semi-norm and the weighted L^2-norm together with the approximation result already shown:

$$
\begin{aligned}
|P_h^{\widetilde{h}} v|_{1,\kappa,\Omega} &\leq C\widetilde{h}^{-1} \|P_h^{\widetilde{h}} v - v\|_{0,\kappa,\Omega} + \|v\|_{1,h,\kappa,\Omega} \leq C\widetilde{h}^{-1} \widetilde{h} \left(\log\left(\frac{2\widetilde{h}}{h} \right) \right)^{1/2} \|v\|_{1,h,\kappa,\Omega} + \|v\|_{1,h,\kappa,\Omega} \\
&\lesssim \left(\log\left(\frac{2\widetilde{h}}{h} \right) \right)^{1/2} \|v\|_{1,h,\kappa,\Omega} \, ,
\end{aligned}
$$

and the proof is complete. □

Appendix B. Proof of (5.19) and (5.18)

Since the results in this section concern only the space V_h^{CR}, we will omit the superscript CR and write A_0 instead of A_0^{CR}, φ_e instead of φ_e^{CR}, etc. We also point out that the two Lemmas that follow are cases of much more general theorem on additive methods which can be found in many texts (see e.g. [77], [69], [67]).

We first prove the identity for the Jacobi method (5.19).

Lemma B.1. *Let R^{-1} be defined via the expression*

$$
R^{-1} A_0 v = \sum_{e \in \mathcal{E}_h^o} \frac{\mathcal{A}_0(v, \varphi_e)}{\mathcal{A}_0(\varphi_e, \varphi_e)} \, \varphi_e .
$$

Then for all $w \in V_h^{CR}$ the following relation holds:

$$
(Rw, w) = \sum_{e \in \mathcal{E}_h^o} \mathcal{A}_0(\varphi_e, \varphi_e) w_e^2 . \tag{B.1}
$$

Proof. Given $w \in V_h^{CR}$, by setting $X = [R^{-1} A_0]^{-1}$ we obtain that

$$
w = R^{-1} A_0 [R^{-1} A_0]^{-1} w = R^{-1} A_0 (Xw) = \sum_{e \in \mathcal{E}_h^o} \frac{\mathcal{A}_0(Xw, \varphi_e)}{\mathcal{A}_0(\varphi_e, \varphi_e)} \, \varphi_e .
$$

In other words, we have $w = \sum_{e \in \mathcal{E}_h^o} w_e \varphi_e$, with $w_e = \frac{\mathcal{A}_0(Xw, \varphi_e)}{\mathcal{A}_0(\varphi_e, \varphi_e)}$. We then easily obtain the desired result by the following, rather straightforward, computations:

$$
\begin{aligned}
\sum_{e \in \mathcal{E}_h^o} \mathcal{A}_0(\varphi_e, \varphi_e) w_e^2 &= \sum_{e \in \mathcal{E}_h^o} \frac{[\mathcal{A}_0(Xw, \varphi_e)]^2}{\mathcal{A}_0(\varphi_e, \varphi_e)} = \sum_{e \in \mathcal{E}_h^o} \frac{\mathcal{A}_0(Xw, \varphi_e)}{\mathcal{A}_0(\varphi_e, \varphi_e)} \mathcal{A}_0(Xw, \varphi_e) \\
&= \sum_{e \in \mathcal{E}_h^o} w_e \mathcal{A}_0(Xw, \varphi_e) = \mathcal{A}_0(Xw, \sum_{e \in \mathcal{E}_h^o} w_e \varphi_e) = \mathcal{A}_0(Xw, w) \\
&= (A_0 [R^{-1} A_0]^{-1} w, w) = (Rw, w).
\end{aligned}
$$

□

The following Lemma verifies the identity for $(B^{-1}v, v)$ given in (5.18).

Lemma B.2. *Let B be defined as in* (5.13)*. We then have*

$$(B^{-1}v, v) = \inf_{\chi \in V_{\widetilde{h}}^{\mathrm{conf}}} [\mathcal{R}(v-\chi, v-\chi) + a(\chi, \chi)].$$

Proof. Let $\chi \in V_{\widetilde{h}}^{\mathrm{conf}}$ be arbitrary and $\eta = (A^C)^{-1}Q^C B^{-1}v$. It is clear that $A^C\eta = Q^C B^{-1}v$ and moreover,

$$v - \eta = v - (A^C)^{-1}Q^C B^{-1}v = BB^{-1}v - P^C A^{-1}B^{-1}v = (B - (A^C)^{-1}Q^C)B^{-1}v = R^{-1}B^{-1}v.$$

Setting $\eta_0 = \chi - \eta$, so that $\chi = \eta + \eta_0$, then shows that

$$\begin{aligned} \mathcal{R}(v-\chi, v-\chi) &= (R(v-\eta-\eta_0), v-\eta-\eta_0) \\ &= (R^{-1}(RB^{-1}v - \eta_0), (RB^{-1}v - \eta_0)) \\ &= (RB^{-1}v, B^{-1}v) - 2(\eta_0, B^{-1}v) + (R^{-1}\eta_0, \eta_0), \end{aligned}$$

and also

$$\begin{aligned} a(\chi, \chi) &= a(\eta + \eta_0), \eta + \eta_0) \\ &= a(\eta, \eta) + 2a(\eta, \eta_0) + a(\eta_0, \eta_0) \\ &= (A^C\eta, \eta) + 2(A^C\eta, \eta_0) + a(\eta_0, \eta_0) \\ &= (Q^C B^{-1}v, (A^C)^{-1}Q^C\eta) + 2(B^{-1}v, \eta_0) + a(\eta_0, \eta_0). \end{aligned}$$

Adding the last two identities and using the fact that $a(\eta_0, \eta_0) \geq 0$ and $(R\eta_0, \eta_0) \geq 0$, and applying the definitions of Q^C and B (in that order) then gives

$$\begin{aligned} \mathcal{R}(v-\chi, v-\chi) + a(\chi, \chi) &\geq (RB^{-1}v, B^{-1}v) + (Q^C B^{-1}v, (A^C)^{-1}Q^C B^{-1}v) \\ &= (B^{-1}v, RB^{-1}v) + (B^{-1}v, (A^C)^{-1}Q^C B^{-1}v) \\ &= (B^{-1}v, \underbrace{(R + (A^C)^{-1}Q^C)}_{B} B^{-1}v) = (B^{-1}v, v) \end{aligned}$$

The proof is complete because χ was arbitrary and moreover, equality holds if and only if $\eta_0 = 0$. □

References

[1] S. Agmon. *Lectures on Elliptic Boundary Value Problems*. Prepared for publication by B. Frank Jones, Jr. with the assistance of George W. Batten, Jr. Van Nostrand Mathematical Studies, No. 2. D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto-London, 1965.

[2] P. F. Antonietti and B. Ayuso. Schwarz domain decomposition preconditioners for discontinuous Galerkin approximations of elliptic problems: non-overlapping case. *Math. Model. Numer. Anal.*, 41(1):21–54, 2007.

[3] P. F. Antonietti and B. Ayuso. Multiplicative Schwarz methods for discontinuous Galerkin approximations of elliptic problems. *Math. Model. Numer. Anal.*, 42(3):443–469, 2008.

[4] P. F. Antonietti and B. Ayuso. Two-level Schwarz preconditioners for super penalty discontinuous Galerkin methods. *Commun. Comput. Phys.*, to appear.

[5] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39(5):1749–1779 (electronic), 2001/02.

[6] O. Axelsson. *Iterative Solution Methods*. Cambridge University Press, Cambridge, 1994.

[7] O. Axelsson. Iteration number for the conjugate gradient method. *Mathematics and Computers in Simulation*, 61(3-6):421–435, 2003. MODELLING 2001 (Pilsen).

[8] B. Ayuso de Dios, F. Brezzi and L. D. Marini. L^2-estimates for the DG IIPG-0 scheme. *Technical report Pavia (2010)*, submitted.

[9] B. Ayuso de Dios and L. Zikatanov. Uniformly convergent iterative methods for discontinuous Galerkin discretizations. *J. Sci. Comput.*, 40(1-3):4–36, 2009.

[10] R. E. Bank, T. Dupont, and H. Yserentant. The hierarchical basis multigrid method. *Numerische Mathematik*, 52:427–458, 1988.

[11] A. Barker, S. Brenner, E.-H. Park, and L.-Y. Sung. Two-level additive Schwarz preconditioners for a weakly over-penalized symmetric interior penalty method. *Journal of Scientific Computing*, pages 1–23, 2010. 10.1007/s10915-010-9419-5.

[12] J. H. Bramble, J. E. Pasciak, and A. H. Schatz. The construction of preconditioners for elliptic problems by substructuring, IV. *Mathematics of Computation*, 53:1–24, 1989.

[13] J. H. Bramble, J. E. Pasciak, and J. Xu. Parallel multilevel preconditioners. *Math. Comp.*, 55(191):1–22, 1990.

[14] J. H. Bramble, J. E. Pasciak, and J. Xu. Parallel multilevel preconditioners. In *Third International Symposium on Domain Decomposition Methods for Partial Differential Equations (Houston, TX, 1989)*, pages 341–357. SIAM, Philadelphia, PA, 1990.

[15] J. H. Bramble and J. Xu. Some estimates for a weighted L^2 projection. *Mathematics of Computation*, 56:463–476, 1991.

[16] A. Brandt, S. F. McCormick, and J. W. Ruge. Algebraic multigrid (AMG) for automatic multigrid solution with application to geodetic computations. Tech. Rep., Institute for Computational Studies, Colorado State Univ ersity, 1982.

[17] S. C. Brenner. Poincaré–Friedrichs inequalities for piecewise H^1 functions. *SIAM Journal on Numerical Analysis*, 41(1):306–324, 2003.

[18] S. C. Brenner, J. Cui, and L.-Y. Sung. Multigrid methods for the symmetric interior penalty method on graded meshes. *Numer. Linear Algebra Appl.*, 16(6):481–501, 2009.

[19] S. C. Brenner and L. Owens. A W-cycle algorithm for a weakly over-penalized interior penalty method. *Comput. Methods Appl. Mech. Engrg.*, 196(37-40):3823–3832, 2007.

[20] S. C. Brenner and J. Zhao. Convergence of multigrid algorithms for interior penalty methods. *Appl. Numer. Anal. Comput. Math.*, 2(1):3–18, 2005.

[21] F. Brezzi, B. Cockburn, L. D. Marini, and E. Süli. Stabilization mechanisms in discontinuous Galerkin finite element methods. *Comput. Methods Appl. Mech. Engrg.*, 195(25-28):3293–3310, 2006.

[22] K. Brix, M. Campos Pinto, and W. Dahmen. A multilevel preconditioner for the interior penalty discontinuous Galerkin method. *SIAM J. Numer. Anal.*, 46(5):2742–2768, 2008.

[23] K. Brix, M. Campos Pinto, W. Dahmen, and R. Massjung. Multilevel preconditioners for the interior penalty discontinuous Galerkin method. II. Quantitative studies. *Commun. Comput. Phys.*, 5(2-4):296–325, 2009.

[24] E. Burman and B. Stamm. Low order discontinuous Galerkin methods for second order elliptic problems. *SIAM J. Numer. Anal.*, 47(1):508–533, 2008.

[25] E. Burman and P. Zunino. A domain decomposition method based on weighted interior penalties for advection-diffusion-reaction problems. *SIAM Journal on Numerical Analysis*, 44(4):1612–1638, 2006.

[26] L. Chen, M. Holst, J. Xu, and Y. Zhu. Local multilevel preconditioners for elliptic equations with jump coefficients on bisection grids. *Arxiv preprint arXiv:1006.3277*, 2010.

[27] D. Cho, J. Xu, and L. Zikatanov. New estimates for the rate of convergence of the method of subspace corrections. *Numerical Mathematics. Theory, Methods and Applications*, 1(1):44–56, 2008.

[28] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. North-Holland Publishing Co., Amsterdam, 1978. Studies in Mathematics and its Applications, Vol. 4.

[29] P. G. Ciarlet. Basic error estimates for elliptic problems. In *Handbook of numerical analysis, Vol. II*, Handb. Numer. Anal., II, pages 17–351. North-Holland, Amsterdam, 1991.

[30] B. Cockburn, O. Dubois, J. Gopalakrishnan, and S. Tan. Multigrid for an HDG method. Submitted, 2010.

[31] R. K. Coomer and I. G. Graham. Massively parallel methods for semiconductor device modelling. *Computing*, 56(1):1–27, 1996.

[32] D. A. Di Pietro, A. Ern, and J.-L. Guermond. Discontinuous Galerkin methods for anisotropic semidefinite diffusion with advection. *SIAM J. Numer. Anal.*, 46(2):805–831, 2008.

[33] V. A. Dobrev. *Preconditioning of discontinuous Galerkin methods for second order elliptic problems*. PhD thesis, Texas A& M University, Department of Mathematics, College Station, TX 77843-3368, December 2007.

[34] V. A. Dobrev, R. D. Lazarov, P. S. Vassilevski, and L. T. Zikatanov. Two-level preconditioning of discontinuous Galerkin approximations of second-order elliptic equations. *Numer. Linear Algebra Appl.*, 13(9):753–770, 2006.

[35] V. A. Dobrev, R. D. Lazarov, and L. T. Zikatanov. Preconditioning of symmetric interior penalty discontinuous Galerkin FEM for elliptic problems. In *Domain decomposition methods in science and engineering XVII*, volume 60 of *Lect. Notes Comput. Sci. Eng.*, pages 33–44. Springer, Berlin, 2008.

[36] V. Dolejší, M. Feistauer, and J. Felcman. On the discrete Friedrichs inequality for nonconforming finite elements. *Numer. Funct. Anal. Optim.*, 20(5-6):437–447, 1999.

[37] M. Dryja. On discontinuous Galerkin methods for elliptic problems with discontinuous coefficients. *Computational Methods in Applied Mathematics*, 3(1):76–85, 2003.

[38] M. Dryja, J. Galvis, and M. Sarkis. BDDC methods for discontinuous Galerkin discretization of elliptic problems. *J. Complexity*, 23(4-6):715–739, 2007.
[39] M. Dryja, J. Galvis, and M. Sarkis. Neumann-Neumann methods for a DG discretization of elliptic problems with discontinuous coefficients on geometrically nonconforming substructures. Technical Report Serie A 634, Instituto de Matematica Pura e Aplicada, Brazil, 2009. submitted.
[40] M. Dryja and M. Sarkis. FETI-DP method for DG discretization of elliptic problems with discontinuous coefficients. Technical report, Instituto de Matematica Pura e Aplicada, Brazil, 2010. submitted.
[41] M. Dryja, M. V. Sarkis, and O. B. Widlund. Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions. *Numerische Mathematik*, 72(3):313–348, 1996.
[42] M. Dryja, B. F. Smith, and O. B. Widlund. Schwarz analysis of iterative substructuring algorithms for elliptic problems in three dimensions. *SIAM J. Numer. Anal.*, 31(6):1662–1694, 1994.
[43] M. Dryja and O. B. Widlund. Schwarz methods of Neumann-Neumann type for three-dimensional elliptic finite element problems. *Comm. Pure Appl. Math.*, 48(2):121–155, 1995.
[44] X. Feng and O. A. Karakashian. Two-level additive Schwarz methods for a discontinuous Galerkin approximation of second order elliptic problems. *SIAM J. Numer. Anal.*, 39(4):1343–1365 (electronic), 2001.
[45] J. Frank and C. Vuik. On the construction of deflation-based preconditioners. *SIAM Journal on Scientific Computing*, 23(2):442–462, 2001.
[46] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, third edition, 1996.
[47] J. Gopalakrishnan and G. Kanschat. A multilevel discontinuous Galerkin method. *Numer. Math.*, 95(3):527–550, 2003.
[48] I. G. Graham and M. J. Hagger. Unstructured additive Schwarz-conjugate gradient method for elliptic problems with highly discontinuous coefficients. *SIAM Journal on Scientific Computing*, 20:2041–2066, 1999.
[49] I. G. Graham and R. Scheichl. Robust domain decomposition algorithms for multiscale PDEs. *Numer. Methods Partial Differential Equations*, 23(4):859–878, 2007.
[50] M. Griebel and P. Oswald. On the abstract theory of additive and multiplicative Schwarz algorithms. *Numer. Math.*, 70(2):163–180, 1995.
[51] W. Hackbusch. *Iterative Solution of Large Sparse Systems of Equations*, volume 95 of *Applied Mathematical Sciences*. Springer-Verlag New York, Inc., 1994.
[52] P. R. Halmos. *Finite-Dimensional Vector Spaces*. Springer-Verlag, New York, second edition, 1974. Undergraduate Texts in Mathematics.
[53] B. Heise and M. Kuhn. Parallel solvers for linear and nonlinear exterior magnetic field problems based upon coupled FE/BE formulations. *Computing*, 56(3):237–258, 1996. International GAMM-Workshop on Multi-level Methods (Meisdorf, 1994).
[54] C. E. Kees, C. T. Miller, E. W. Jenkins, and C. T. Kelley. Versatile two-level Schwarz preconditioners for multiphase flow. *Comput. Geosci.*, 7(2):91–114, 2003.
[55] A. Klawonn, O. Widlund, and M. Dryja. Dual-primal FETI methods for three-dimensional elliptic problems with heterogeneous coefficients. *SIAM J. Numer. Anal.*, 40(1):159–179, 2002.
[56] J. K. Kraus and S. K. Tomar. A multilevel method for discontinuous Galerkin approximation of three-dimensional anisotropic elliptic problems. *Numer. Linear Algebra Appl.*, 15(5):417–438, 2008.
[57] J. K. Kraus and S. K. Tomar. Multilevel preconditioning of two-dimensional elliptic problems discretized by a class of discontinuous Galerkin methods. *SIAM J. Sci. Comput.*, 30(2):684–706, 2008.
[58] J. Mandel and M. Brezina. Balancing domain decomposition for problems with large jumps in coefficients. *Math. Comp.*, 65(216):1387–1401, 1996.
[59] J. Meza and R. Tuminaro. A multigrid preconditioner for the semiconductor equations. *SIAM Journal on Scientific Computing*, 17:118–132, 1996.
[60] P. Oswald. *Multilevel Finite Element Approximation, Theory and Applications*. Teubner Skripten zur Numerik. Teubner Verlag, Stuttgart, 1994.
[61] P. Oswald. On the robustness of the BPX-preconditioner with respect to jumps in the coefficients. *Mathematics of Computation*, 68:633–650, 1999.
[62] F. Prill, M. Lukáčová-Medvid̆ová, and R. Hartmann. Smoothed aggregation multigrid for the discontinuous Galerkin method. *SIAM J. Sci. Comput.*, 31(5):3503–3528, 2009.
[63] Y. Saad. *Iterative Methods for Sparse Linear Systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, second edition, 2003.
[64] M. Sarkis. Multilevel methods for P_1 nonconforming finite elements and discontinuous coefficients in three dimensions. In *Domain decomposition methods in scientific and engineering computing (University Park, PA, 1993)*, volume 180 of *Contemp. Math.*, pages 119–124. Amer. Math. Soc., Providence, RI, 1994.
[65] M. Sarkis. Nonstandard coarse spaces and Schwarz methods for elliptic problems with discontinuous coefficients using non-conforming elements. *Numer. Math.*, 77(3):383–406, 1997.

[66] R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Mathematics of Computation*, 54:483–493, 1990.
[67] B. F. Smith, P. E. Bjørstad, and W. D. Gropp. *Domain Decomposition*. Cambridge University Press, Cambridge, 1996. Parallel multilevel methods for elliptic partial differential equations.
[68] R. Stenberg. Mortaring by a method of J. A. Nitsche. In *Computational mechanics (Buenos Aires, 1998)*, pages CD–ROM file. Centro Internac. Métodos Numér. Ing., Barcelona, 1998.
[69] A. Toselli and O. Widlund. *Domain Decomposition Methods: Algorithms and Theory*. Springer Series in Computational Mathematics, 2005.
[70] P. S. Vassilevski. On two ways of stabilizing the hierarchical basis multilevel methods. *SIAM Review*, 39(1):18–53, 1997.
[71] P. S. Vassilevski. *Multilevel Block Factorization Preconditioners*. Springer-Verlag, 2008.
[72] C. Vuik, A. Segal, and J. A. Meijerink. An efficient preconditioned CG method for the solution of a class of layered problems with extreme contrasts in the coefficients. *Journal of Computational Physics*, 152(1):385–403, June 1999.
[73] W. L. Wan, T. F. Chan, and B. Smith. An energy-minimizing interpolation for robust multigrid methods. *SIAM Journal on Scientific Computing*, 21(4):1632–1649 (electronic), 2000.
[74] C. Wang. Fundamental models for fuel cell engineering. *an invited review article for Chemical Reviews*, 104:4727–4766, 2004.
[75] J. Wang. New convergence estimates for multilevel algorithms for finite-element approximations. *Journal of Computational and Applied Mathematics*, 50:593–604, 1994.
[76] Z. Wang, C. Wang, and K. Chen. Two-phase flow and transport in the air cathode of proton exchange membrane fuel cells. *J.Power Sources*, 94:40–50, 2001.
[77] J. Xu. Iterative methods by space decomposition and subspace correction. *SIAM Rev.*, 34(4):581–613, 1992.
[78] J. Xu and Y. Zhu. Uniform convergent multigrid methods for elliptic problems with strongly discontinuous coefficients. *Mathematical Models and Methods in Applied Science*, 18(1):77 –105, 2008.
[79] J. Xu and J. Zou. Some nonoverlapping domain decomposition methods. *SIAM Rev.*, 40(4):857–914, 1998.
[80] Y. Zhu. Domain decomposition preconditioners for elliptic equations with jump coefficients. *Numerical Linear Algebra with Applications*, 15(2-3):271–289, 2008.
[81] L. Zikatanov. Two-sided bounds on the convergence rate of two-level methods. *Numerical Linear Algebra with Applications*, 15(5):439 – 454, 2008.

E-mail address: `bayuso@crm.cat`

E-mail address: `mholst@math.ucsd.edu`

E-mail address: `zhu@math.ucsd.edu`

E-mail address: `ltz@math.psu.edu`